THE PROCEEDINGS OF
THE PHYSICAL SOCIETY
Section A

VOL. 62, PART 8 1 August 1949 No. 356 A

Steam in the Ring Discharge

By G. I. FINCH*

Presidential Address, delivered 6th May 1949

ABSTRACT. The behaviour of steam and its decomposition products in the ring discharge has been examined.

Dry hydrogen is not dissociated. The production of atomic hydrogen is dependent upon the presence of steam which dissociates into hydroxyl and atomic hydrogen. A secondary source of atomic hydrogen is then afforded by the interaction of hydroxyl with molecular hydrogen.

The escape from the discharge of atomic hydrogen, a long-lived species, favours the dissociation of steam. Mercury vapour, on the other hand, inhibits the formation of atomic hydrogen and thus leads to a high equilibrium steam concentration.

Unlike dry hydrogen, dry oxygen is dissociated into atoms, but these have a short life as such and recombine in the discharge to form molecular oxygen and ozone.

The reaction mechanisms occurring in the discharge are discussed in the light of spectrographic results.

§1. INTRODUCTION

BETWEEN 1926 and 1935 my collaborators and I carried out a systematic study of the combustion of hydrogen and carbon monoxide by oxygen in the cathode glow of the electric discharge maintained between metallic electrodes. Our object was to elucidate the nature of the mechanisms of these combustion processes, and the results and conclusions drawn therefrom were summarized in the *Journal of the Chemical Society* (Finch 1935). It may be remarked that, during the progress of this work, it became apparent that the results were also of practical interest, in that they threw fresh light on the chemistry of the thermionic valve, particularly in relation to the electron emissivity of the cathode.

Whilst certain phenomena arising out of the use of metallic electrodes in our experiments had helped towards an understanding of some aspects of the general mechanism of combustion, it became evident that, in other directions, they tended to obscure the issues. Indeed, as regards the undoubtedly complex mechanism of the homogeneous combustion of hydrogen by oxygen to steam, we could, in 1935, say no more than that, when the combustible mixture was rigidly dried, the initial step was the formation of the hydroxyl radical and that, in the undried mixture, steam was first dissociated into atomic hydrogen and hydroxyl, whereupon the hydroxyl somehow acted as a powerful promoter in the succeeding stages, leading ultimately to the formation of steam. No positive rôle in the combustion process could be assigned to atomic hydrogen. This was remarkable since, under the conditions of our experiments, atomic hydrogen was known to be produced in considerable quantities. Accordingly it was decided to examine the equilibrium between steam and its dissociation products in the electrodeless discharge under conditions which would eliminate surface effects, or at least permit of an adequate appreciation of the part played by them.

---

* The Presidential Address will also appear in *Proc. Phys. Soc.* for September 1949, Section B.

§ 2. APPARATUS AND GENERAL PROCEDURE

The gas (steam, $2H_2+O_2$, H_2, O_2 or rare gas, or various mixtures thereof) was contained in the closed circulation system shown in Figure 1. This was assembled from the cleaned components using a dry, filtered, glass-blowing air supply, and was so constructed that the circulating gas came into contact only with glass, mercury or quartz, thus avoiding the contaminations introduced by lubricated stopcocks and rubber joints which have in the past so often vitiated the results of work on the combustion of hydrogen. The glass and quartz components were cleaned with chromic acid and distilled water ; the mercury, previously washed in the usual manner and distilled, was redistilled into the appropriate vessels ; and the phosphoric oxide was redistilled in oxygen until free from the lower oxides (Finch and Fraser 1926). The major remaining sources of impurity (other than mercury vapour) were the gases, chiefly carbon monoxide and nitrogen, dissolved in the glass of the discharge vessel. As will be seen later, getting rid of these gases presented no difficulty.

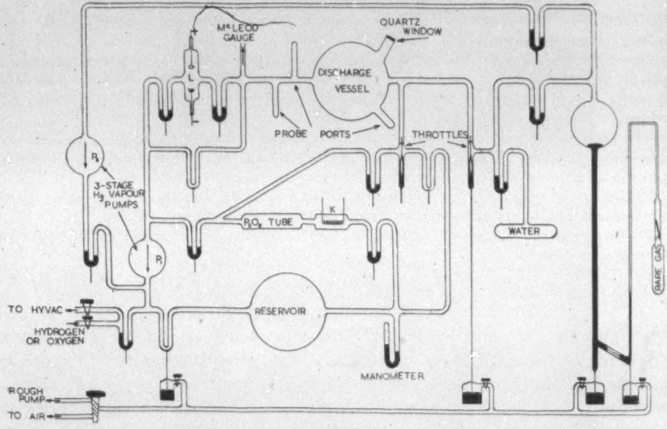


Figure 1.

In Figure 1 the apparatus is shown with the barometric cut-offs and throttles set for the passage of the gas through the main circulation system consisting of the discharge vessel (10 cm. diameter), the three-stage mercury vapour pump, P_1, and the reservoir (10 cm. diameter). In L a high-tension arc could be maintained between a nickel anode and a molten lithium cathode for the purification of the rare gas admitted to the system. By a suitable setting of the appropriate cut-offs, the vessel K and the phosphoric oxide tube could be inserted into the circulation system between the reservoir and discharge vessels for the purpose of the quantitative analysis of steam-hydrogen-oxygen mixtures. Thus, with the platinum wire coil in K cold, the steam was removed with a corresponding change in the manometer level. Next, with the coil heated, hydrogen and oxygen in their combining proportions were removed, with a further change in the manometer level, the remaining gas being excess hydrogen, oxygen or known added inert diluent such as, for example, neon. The gaseous contents of the discharge vessel–reservoir circulation system could be well dried by circulation through the interposed phosphoric oxide tube, though that degree of rigid drying which exerts a marked effect in increasing the resistance of hydrogen to ignition could not be achieved. It is well known to workers in this field that, despite all precautions, the use of mercury seals and cut-offs invariably introduces traces of moisture.

The normal working pressure of gas in the reservoir during circulation was between about 7 and 10 mm., whilst that in the discharge vessel was maintained at between 0·026 and 0·03 mm. Hg by means of the pump, P_1, and the appropriate throttle. The rate of circulation through the discharge vessel was rather over 200 cm^3 per second ; this was measured in a manner which will be obvious from Figure 1. Owing to the fall in pressure at the throttle, the partial pressure of mercury vapour in the discharge vessel could not exceed 3×10^{-6} mm., the rate of circulation being more than sufficient to prevent diffusion of mercury from the McLeod gauge, cut-offs or pump towards the discharge vessel.

Short sealed-off tubes were fused, as shown, into the discharge vessel and the 1·2 cm. bore tube, nearly 1 metre long, between the discharge vessel and the McLeod gauge entry tube. These tubes served for the sealing in of probes projecting into the discharge or gas stream. Three probe tubes, two of which are shown in Figure 1, were situated at 5, 20 and 50 cm. respectively beyond the discharge vessel outlet. In all other respects Figure 1 is self-explanatory.

The discharge was maintained by means of a high-frequency current in a solenoid, 13 cm. in length and internal diameter, mounted coaxially with the discharge vessel, and consisting of seven turns of 4 mm. diameter copper tubing. The current was obtained from a valve (Mullard S.W.9) oscillator in a modified Hartley circuit. The valve filament was run at its rated value and, unless otherwise stated, the input to the oscillator was 160 watts. The frequency was 26 Mc/s. as measured by a Southworth-Lecher (Southworth 1920) wire system. Further, the gas pressure in the discharge vessel was maintained at between about 0·026 and 0·03 mm. Hg, preferably nearer the former, the McLeod gauge readings being corrected for the experimentally determined equilibrium steam contents when moist gaseous mixtures were in use.

Under these conditions the discharge in steam consisted of a bright rose-pink annulus coaxial with the discharge vessel, from the cool walls of which it was separated by a dark sheath about 2 or 3 mm. thick. In cross section the diameter of the bright core of the annulus was about 4 cm., the intensity of the glow decreasing considerably near the centre of the discharge tube. It will be evident that this was a ring discharge, the theory of which was outlined in the 13th Guthrie Lecture by J. J. Thomson (1928). In such a discharge the currents form closed circuits in the gas, though it is possible that, under the conditions of these experiments, the actual initiation of the discharge was effected by the electrostatic field across the ends of the solenoid (Mackinnon 1929, Esclangen 1934). It will be appreciated that conditions in this ring discharge closely resemble those in the positive column.

The evacuated discharge vessel was finally cleansed and gases absorbed in its walls sufficiently removed, first by prolonged baking at about 350° C. and then by maintaining the discharge in a stream of steam drawn from the vessel containing CO_2-free water, past the appropriate throttle, through the cool discharge vessel and the pump P_1 backed by the 'Hyvac' pump. This also served to free the water in its storage vessel from dissolved air. After evacuation following upon such cleansing, the main circulation route was restored and the system filled with the appropriate gas to the desired pressures in the discharge and reservoir vessels.

A mixture of hydrogen and oxygen in combining proportions was frequently used in the experiments to be described below. It was conveniently prepared in a high state of purity by drawing pure steam through the discharge and drying the products accumulated in the reservoir by subsequent circulation through the phosphoric oxide tube. Oxygen was obtained by heating recrystallized potassium permanganate and then washing with a caustic potash solution. Hydrogen was prepared by electrolysis of a saturated recrystallized baryta solution and, when necessary, was subsequently freed from oxygen and hydrogen oxides by circulation past the heated platinum spiral K and through the phosphoric oxide tube.

§ 3. EXPERIMENTAL RESULTS AND CONCLUSIONS

Series 1—*Dry and moist hydrogen.* A puzzling feature of our work on the cathodic combustion of hydrogen was that, as has already been remarked, no evidence was found to justify the assigning of a direct rôle in combustion to the atomic hydrogen, which was known to be produced in considerable quantities. Accordingly it was decided to examine first the behaviour of hydrogen, dried as well as possible by circulation over phosphoric oxide, in the electrodeless discharge.

After a preliminary drying of the discharge tube walls by means of a prolonged circulation of dry hydrogen with the phosphoric oxide tube in the circuit, the filament K glowing and the discharge in action, the apparatus was evacuated and charged with dried hydrogen. It was then found that the hydrogen could be circulated through the discharge for long periods without any change in pressure and that, in the early stages of circulation, the spectrum (Figure 2*) of the bluish-white discharge consisted mainly of the secondary spectrum due to neutral molecular hydrogen. The hydroxyl bands (2808 and 3064 A.) and the more intense lines (H_α 6563, H_β 4861 and H_γ 4350 A.) of the Balmer series, due to neutral atomic hydrogen, were visible but, with the exception of H_α, were feeble compared with those of the secondary series. Also in evidence were a few mercury lines, in particular 2536, 3650 and 4358 A. With increasing time of circulation and drying, during which precautions were taken to avoid shaking or otherwise disturbing the mercury in the McLeod gauge and the cut-offs associated with the main circulation system, the degree of dryness of the gas evidently improved until, except for the brilliant secondary series and the few mercury lines mentioned above, only the 3064 A. hydroxyl band-head region and the red H_α line still remained visible. Thus it seems likely that, had it been possible to achieve a rigid drying of the hydrogen, the Balmer series and hydroxyl bands would have been eliminated.

---

* For Figures 2 to 6 see Plates.

These results confirm, in all essential respects, the experimental findings of Wood (1920, 1921, 1922 a, b), Copeland (1930), Poole (1937) and other workers. Wood concluded that atomic hydrogen was copiously formed in the discharge passing through dry hydrogen and attributed the virtual non-appearance of the Balmer series to a catalytic recombination of the atomic hydrogen to the molecular form at the dry walls of the vessel. Other workers have since accepted this hypothesis.

For the three following reasons, however, I have not been able to subscribe to this view without submitting it to the test of further experiment. In the first place, Wood and others worked with gas pressures of the order of 1 mm. ; it seems improbable, therefore, that any considerable proportion of hydrogen atoms formed in the discharge could arrive at the walls of the vessel without being excited and radiating, and thus giving rise to the Balmer series. Next, the catalytic inertness of cool, clean dry glass is well known, particularly to workers in the field of gaseous reactions in closed vessels. Thus Strutt (1911) found that the greenish glow of active nitrogen in a clean glass vessel persisted for several minutes after excitation, although it was extinguished almost instantaneously on introducing into the vessel a superfically oxidized copper wire. Finally, although the quantum theory indicates the possibility of the dissociation of the hydrogen molecule by electron impact, the necessary excitation energy would be 18 ev. (Mott and Massey 1933), and it is doubtful if such high energies were available in the ring discharge maintained under the conditions set forth above. The matter is one which was put to the test of crucial experiment as follows.

The pointed end of a glass needle was wetted to a length of about 3 mm. with diluted waterglass, dusted with yellow mercuric oxide powder, dried *in vacuo* and sealed into the discharge vessel probe port. The tip of the needle was about level with the contour of the discharge vessel so that, if any hydrogen atoms were formed in the discharge, some would reach the mercuric oxide-coated needle tip directly. The circulation system with the phosphoric oxide tube in circuit and the wire K at red heat was then filled with hydrogen which was circulated for 26 hours. The discharge was then started in the, by this time, relatively well-dried, oxygen-free hydrogen, and the circulation continued. At first no change could be seen in the appearance of the mercuric oxide ; it was not until after 20 minutes that the first signs of reduction, consisting in a blackening due to the formation of mercurous oxide, could be detected. The drying tube was then by-passed out of the circulation system and a little oxygen admitted to the reservoir contents, whereupon, within a few seconds, the yellow changed to black oxide and this, in turn, was quickly further reduced to minute droplets of mercury which slowly evaporated. In another experiment a similar procedure was followed, except that instead of oxygen about 2 per cent by volume of steam was added to the reservoir vessel contents. Here again the same remarkable acceleration in the rate of reduction of the mercuric oxide occurred. As will be shown later, the agent responsible for the reduction of both mercury oxides was atomic hydrogen.

It is difficult to resist drawing from these experimental results the conclusion that, apart from traces due to imperfect drying and revealed spectrographically, atomic hydrogen is not formed when an electrodeless discharge passes through dry hydrogen. Hence, the hitherto generally accepted view that the hydrogen molecule is atomized by the discharge in dry hydrogen and that the absence of the Balmer spectrum is to be accounted for by a recombination to the molecular state at the surface of the discharge vessel is untenable. These experiments, of course, afford no information as to the relative activities of dry and moist glass surfaces in causing the recombination of hydrogen atoms to molecules ; but they do show that dry hydrogen is not dissociated by the discharge, and that the minute amount of atomic hydrogen formed is attributable to that trace of moisture which can hardly be eliminated in an experiment of this kind carried out in an apparatus equipped with barometric mercury cut-offs.

Series 2—*The steam equilibrium.* Preliminary experiments in which steam had been circulated through the discharge, primarily with a view to outgassing the discharge vessel, had shown that much of the steam was decomposed and that the only other products received into the reservoir vessel consisted of hydrogen and oxygen in their combining proportions. In further such experiments steam was drawn through the discharge, and the products accumulated to a pressure of about 10 mm. in the previously evacuated reservoir, the cut-off between this and the discharge vessel being closed. On interrupting the steam supply and restoring the main circuit, no further change in pressure occurred, even after several hours' circulation. Analyses immediately before resumption of, or after, prolonged circulation revealed no change in the steam concentration, the remaining products always consisting of hydrogen and oxygen in their combining proportions. It is clear, therefore, that the steam dissociation equilibrium was attained in a single passage of the steam through the discharge and, further, that no gas was lost by, for example, combination with mercury.

In the course of these and other exploratory experiments it was also found that the extent of this dissociation was much affected by the state of the surface of the walls of the discharge vessel. For example, with clean walls and starting with either steam or the equivalent mixture of hydrogen and oxygen, the products at equilibrium contained between 16·5 and 19 per cent by volume of steam. On repeating such an experiment, but with a fine (37 S.W.G.) copper wire protruding into the discharge through the appropriate port, the wire glowed and the equilibrium steam concentration rose progressively in the course of five two-hour runs to between 38 and 40 per cent. By this time the walls of the discharge vessel near the wire were faintly discoloured, presumably by evaporated copper which may or may not have been oxidized. The glowing of the wire was due to some energetic surface reaction, and not merely to eddy-current heating, as was shown by the fact that on reducing the pressure in the discharge vessel, by closing the throttle and thus extinguishing the gas discharge, the wire at once ceased to glow. The wire was then removed and the port sealed off, whereupon further runs, in which equilibrium was approached sometimes from one side and sometimes from the other, continued to yield equilibrium steam concentrations approaching 40 per cent as before. Thus there was no tendency for this concentration to increase, as had been the case when the copper wire protruded into the discharge.

The discharge vessel was now cut out of the apparatus, cleaned with fuming nitric acid followed by chromic acid, washed with distilled water and finally dried with warm filtered air. The vessel was then re-sealed into the circulation apparatus, using dry filtered air for glass-blowing purposes. It was then found that this cleaning had restored the discharge vessel to its original state, equilibrium steam concentrations of about 18 per cent being once more obtained.

These results point to a promotion, by the copper stain on the discharge vessel walls, of some surface reaction or reactions which, either directly or through the products formed, favoured the formation of steam.

The reactants concerned must be relatively long-lived, because they survive passage from the discharge region to the vessel walls across the intervening discharge-free dark zone. Further, bearing in mind the evidence afforded by the glowing copper wire of the highly energetic nature of the reaction, and the testimony of the mercuric oxide probe experiments with moist hydrogen to the arrival of atomic hydrogen at the discharge vessel walls, it seems likely that the surface reaction is the recombination of atomic to molecular hydrogen, and that this may be, indeed, the dominating reaction increasing the equilibrium steam concentration. In other words, the formation of atomic hydrogen seems, somehow or other, to hinder the formation of steam.

Further light was thrown on these matters by the spectrograms (Figure 3) of the discharge in steam and its decomposition products. These are striking in that, apart from the relatively weak mercury spectrum, they consist of the Balmer series (3771, 3797, 3835, 3890, 3970, 4102, 4340, 4861 and 6562 A.), the hydroxyl bands with heads at 2808 and 3064 A., and the rather weak lines, 3947 and 4368 A., due to atomic oxygen. The secondary molecular hydrogen spectrum is absent. Further, the presence or otherwise of copper deposited on the discharge vessel walls had no appreciable effect on the appearance of the spectrograms or on the relative intensities of the hydroxyl band and Balmer spectra, despite the profound effect exerted by the copper stain in the equilibrium steam concentration.

Wood and other workers have obtained spectra under conditions comparable to those outlined above, except that, instead of pure steam, they used moist hydrogen, the steam concentration being in general about $2\frac{1}{2}$ per cent. Apart from the absence of all atomic oxygen lines, their spectra are essentially similar to those obtained in the present experiments. Thus, while Wood and others have shown that copious supplies of long-lived atomic hydrogen are obtained from the discharge in moist hydrogen, the above experiments demonstrate that atomic hydrogen is also freely produced when steam alone is fed to the discharge.

Series 3—*Steam–hydrogen mixtures*. In order to gain further insight into the mechanism of the formation of atomic hydrogen, experiments were now carried out in which a wide range of mixtures of hydrogen with steam, or with hydrogen and oxygen in their combining proportions, were circulated through the discharge. Before each such experiment, however, the circulation system was prepared in the following manner. Glass needles, the points of which had been coated with ash (mainly tin oxide and lime) by calcining a thin smear of sealing wax, were sealed into the three ports with their tips projecting into the tube between the discharge vessel and the McLeod gauge. Another ash-coated probe was sealed into the discharge tube port with its tip level with the contour of the vessel. Dry oxygen was then circulated under standard pressure conditions through the discharge. The reaction product reacted with mercury vapour to deposit in a few minutes broad rings or, rather, bands of yellow mercuric oxide about the openings leading to the McLeod gauge and the rare gas purifier, and just above the entry to the pump, P_1. The apparatus was then evacuated and the circulation system, with the phosphoric oxide tube in circuit, swept out and finally charged with roughly dried hydrogen, which was then still further dried by prolonged circulation during which the platinum spiral, K, was glowing. Throughout these operations the mercuric oxide films remained unchanged. The discharge was now started and spectrographed. Apart from the strong secondary molecular spectrum and the usual lines due to traces of mercury, only the H_α (moderate to weak), H_β (weak) and H_γ (very weak) lines of the Balmer series and the hydroxyl 3064 A. band-head region were visible. The circulating gas had no visible effect on any of the probes coated with sealing-wax ash ; likewise, the mercuric oxide films near the McLeod gauge and the pump were unchanged, even after two hours' circulation.

The apparatus was now evacuated, and the normal circulation route restored and filled to the usual pressures with the required mixture of hydrogen and steam. After a sufficiently prolonged circulation to ensure thorough mixing of the gases, the discharge was started. The range of mixtures examined extended from steam alone to steam (or the equivalent hydrogen–oxygen mixture) diluted with up to 95 per cent of hydrogen. Spectrograms were obtained in each case under uniform conditions of exposure and development. The sealing-wax ash probe inserted through the discharge vessel port was withdrawn after the first experiment and was not replaced.

The results of this series are illuminating. In the first experiment, carried out with steam alone, the mercuric oxide films, including that near the mercury vapour pump inlet, about 1½ metres beyond the discharge vessel, were completely reduced within a few minutes of starting the discharge. The progress of such reduction could, however, be immediately arrested by interrupting the discharge. Further, the particles of ash on the discharge vessel probe and on the two probes nearest to the discharge vessel flashed into incandescence immediately the discharge was started, but ceased to glow when this was interrupted. From time to time glowing fragments of ash would fly off the probes, particularly that in the discharge vessel, and these continued to incandesce after lodging on the wall of the vessel or in the tube beyond it. Indeed, owing to this stripping-off of the ash, the section of tubing carrying the first two ports beyond the discharge vessel had to be removed for cleaning after each experiment and fresh ash-coated probes inserted in preparation for the next. As stated above, the discharge vessel probe was removed and not replaced after the first experiment.

The incandescence of the ash on the second probe beyond the discharge vessel was less brilliant than that of the first probe and, in the case of the third probe some 50 cm. beyond the discharge vessel, an ash particle would sometimes glow, but only in mixtures containing initially between about 30 and 60 per cent of hydrogen.

The spectrograms were similar to those obtained with pure steam (Figure 3), except that the atomic oxygen lines, 3947 and 4368 A., seemed to weaken with increasing hydrogen concentration ; they no longer appeared in a spectrogram from the discharge in a mixture of steam and hydrogen in equal volumes. Visual comparison of the spectrograms, more particularly of the hydroxyl band 2808 A. with H 3797 and H 3835 A., and of the hydroxyl band 3064 A. with H 3890 and H 3970 A., revealed no change with increasing hydrogen concentration in the relative intensities of the hydroxyl band and Balmer series spectra.

Finally, it was found that the steam equilibrium constant decreased with increasing dilution of the mixture by hydrogen. For example, with steam alone the constant $K=(H_2O)^2/[(H_2)^2.(O_2)]=0{\cdot}46$, but was 0·38 and 0·25 in the case of mixtures diluted with 55 and 69 per cent of hydrogen respectively. Thus,

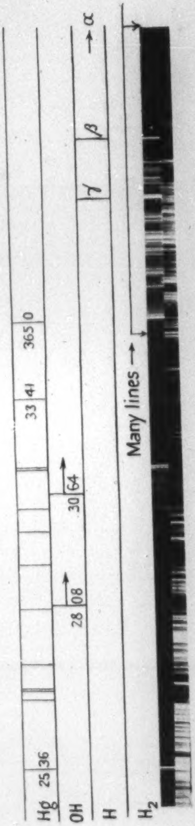


Figure 2. Roughly dried hydrogen.

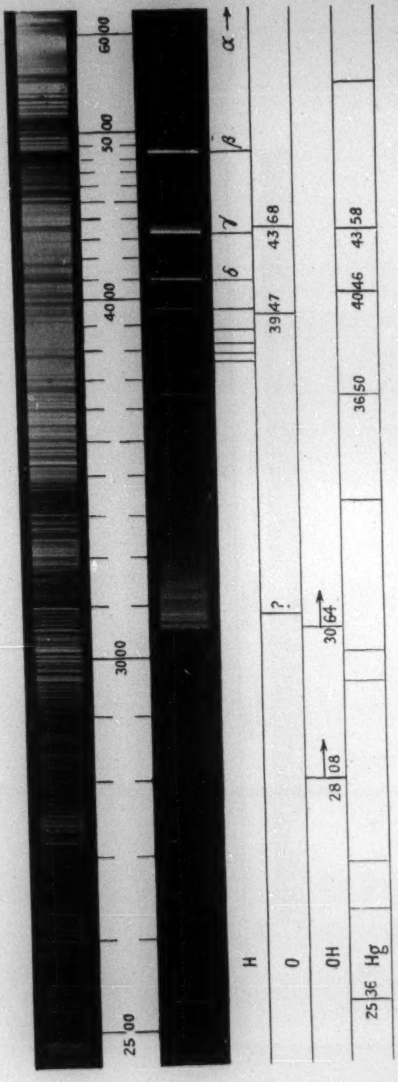


Figure 3. Steam.

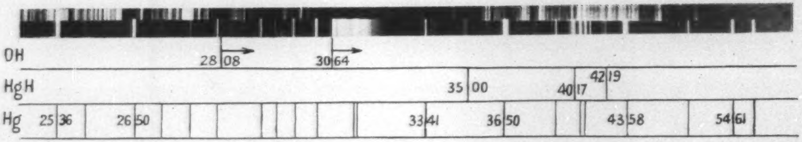


Figure 4. Steam and mercury vapour.

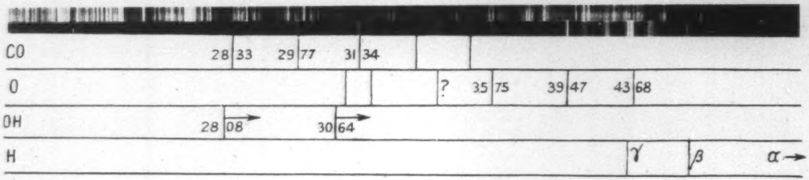


Figure 5. Roughly dried oxygen.

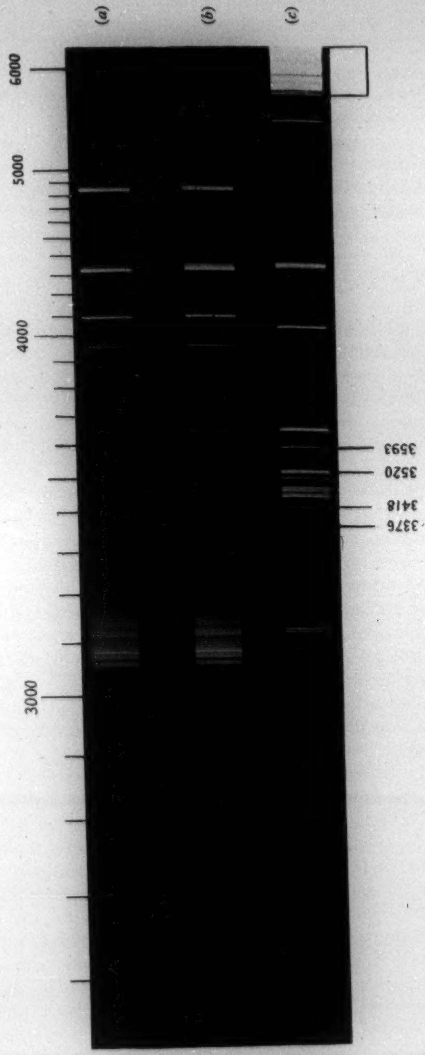


Figure 6. (*a*) Steam; (*b*) 50% steam and 50% neon; (*c*) neon.

although dilution with hydrogen did, in fact, greatly reduce the dissociation of the steam by the discharge, it did not do so to the extent required by the law of mass action.

The reason is not far to seek. The results of these experiments as a whole now afford overwhelming evidence that the first step towards equilibrium in steam or in a steam–hydrogen mixture is the direct dissociation of the steam into hydroxyl and atomic hydrogen. The molecular hydrogen as such is not dissociated by the discharge and, when steam is present in adequate supply (which may possibly be considerably less than 5 per cent), either the hydrogen is not excited or it reacts in some way before it can radiate. This dissociation of steam must be practically irreversible in the gaseous phase, because quantum energy considerations forbid the reverse reaction by which atomic hydrogen and hydroxyl might recombine to form steam, except as a result of necessarily infrequent three-body collisions. Similar considerations apply in the case of the union of hydrogen atoms to the molecular state. On the other hand, the life of the hydroxyl radical (Bonhoeffer and Pearson 1931, Geib and Harteck 1934, Oldenberg 1937) can be terminated by two-body collision processes such as, for example, $2OH = H_2O + O$, and $OH + H_2 = H_2O + H$. Thus the fall in the steam dissociation equilibrium constant with increasing dilution by hydrogen is to be accounted for by the practically irreversible formation, and hence escape, of atomic hydrogen from the reaction zone.

There is evidence from the above experiments that the hydroxyl reacted in the discharge in both the ways suggested. That the first reaction occurred is supported by the appearance of oxygen atom lines in the spectrograms from the discharge in steam, and by their weakening and final disappearance on progressive dilution with hydrogen. As for the second reaction, in which hydroxyl interacts with hydrogen, the likelihood of its occurrence is supported by the following facts and considerations. In the first place, the relative intensities of the hydroxyl and Balmer spectra (Figure 3) were independent of hydrogen dilution of steam over a wide range. This strongly supports the previous conclusion, that the sole primary source of atomic hydrogen is the direct dissociation of steam in the discharge. Next, the fact that the third sealing-wax ash probe glowed only in mixtures of steam diluted by a roughly equal proportion of hydrogen shows that the yield of atomic hydrogen was increased by such dilution. Hence the diluent hydrogen must itself contribute in some way to the ultimate yield of atomic hydrogen. Since, as has already been shown, atomic hydrogen cannot have been produced by direct dissociation of molecular hydrogen, the probable source of this contribution lies in the interaction in the discharge of hydroxyl and molecular hydrogen, whereby atomic hydrogen and steam are formed. This view received strong support from the results, which showed that the recombination of atomic hydrogen at the copper-stained walls of the discharge vessel increased the steam equilibrium concentration by returning molecular hydrogen to the discharge and thus reducing the amount escaping in the atomic state from the reaction zone. The possibility of a surface reaction between hydroxyl and atomic hydrogen in similarly affecting the steam equilibrium will be considered later.

The results of this series of experiments furnish ample evidence of the significance to be attached to the probe and mercuric oxide reactions. Thus, when the discharge was fed with steam, steam diluted with hydrogen, or the equivalent hydrogen–oxygen mixtures, the mercuric oxide films were rapidly reduced, and the sealing-wax ash probes incandesced with a brilliancy which increased as the discharge vessel was approached. Further, all these phenomena ceased immediately the discharge was interrupted, irrespective of whether the circulating gas was steam (or the equivalent hydrogen–oxygen mixture) alone or diluted with hydrogen. Therefore the agent reducing the mercuric oxide could not have been molecular hydrogen, nor could the reaction catalysed by the ash have been the union of hydrogen and oxygen to steam. In addition to the normal equilibrium products entering the reservoir vessel, the only other products for which there is evidence of formation in the discharge are atomic hydrogen, hydroxyl and hydrogen peroxide (Finch and Mahler 1931). Atomic oxygen can be excluded because, although revealed spectrographically in the case of steam traversing the discharge, it was eliminated by sufficient dilution of the reaction mixture with hydrogen. Hence, whereas the reduction of the mercuric oxide films can only be attributed to atomic hydrogen, the incandescence of the ash-coated probes might be ascribed either to the formation of molecular hydrogen or to the interaction of atomic hydrogen with hydroxyl to form steam. The heat evolutions of these reactions are similar, being about 103,000 kcal/gm. mol (see Steacie 1946). Whilst there is no doubt about the longevity of atomic hydrogen, the situation in regard to the life of hydroxyl, though known to be much shorter than that of atomic hydrogen, is less clear. Thus, while Bonhoeffer and Pearson (1931) were unable to find the radical in the gas flowing from a discharge through water vapour, Frost and Oldenberg (1936), using a 21-foot grating, detected hydroxyl for as long as $\frac{1}{8}$ second after their discharge in steam was interrupted. Thermal conditions may, however, have contributed to this result. The reduction of the mercuric oxide probe in the discharge vessel port, when the discharge took place in moist hydrogen, would be difficult to explain if hydroxyl arrived at the probe, and hence also at the surface of the vessel, at a rate comparable with that of the atomic hydrogen. Thus the evidence so far supports the view that the incandescence of the ash-coated probes was due to the recombination of atomic hydrogen. This issue was further tested in the next series of experiments.

Of the three ash-coated probes, all of which were much nearer to the discharge vessel than were the mercuric oxide films, the first two incandesced regularly, whilst the third did so spasmodically, and then only under the most favourable conditions. It would seem that, whilst the reduction of the mercuric oxide film formed a sensitive test for the presence of atomic hydrogen, the sealing-wax ash probes could respond only to a relatively high concentration.

Series 4—*Further probe experiments.* In the first experiment of this series a closely wound spiral nearly 1 cm. in diameter, of 28 S.W.G. copper wire, was inserted into the exhaust tube leading from the discharge vessel. The spiral, previously cleaned and reduced by hot hydrogen, extended from near the second port, which was fitted with a mercuric oxide-coated probe, to within 0·5 cm. from the discharge vessel. The equilibrium steam concentration was then determined, starting with steam or the equivalent hydrogen–oxygen mixture. The mean steam yield from eight such runs was 19·2 per cent, a value which did not differ significantly from that (18·9 per cent) afforded by a similar series of runs carried out before inserting the spiral. The presence of the copper spiral, however, practically inhibited the reduction of the mercuric oxide, thus showing that the bulk of the atomic hydrogen leaving the discharge vessel was, on passing through the spiral, transformed to the molecular state. This, together with the fact that the steam yield was virtually unaffected by the presence or otherwise of the spiral, proves conclusively that hydroxyl could not have survived long enough after leaving the discharge to reach the spiral in appreciable quantities ; otherwise the equilibrium steam concentration must have increased, which it did not do. Hence these results confirm that the ash incandescence observed in the previous experiments was due solely to recombination of atomic hydrogen.

The spiral was now removed and the corresponding tube section cleaned and re-fused into the apparatus. The glass pinch of a tungsten filament lamp was then sealed into the first probe port beyond the discharge vessel, the lamp filament having been replaced by a 2·5 cm. long fine (37 S.W.G.) platinum wire and the current-carrying posts shrouded with quill tubing. The wire projected in the form of a hoop into the path of the gas stream from the discharge vessel.

In the next experiment a mixture of hydrogen and oxygen was circulated under normal conditions but without the discharge, the phosphoric oxide tube being included in the circulation path. After drying the gas in this manner until the manometer readings were steady, the platinum filament was brought to a bright red heat and circulation continued. The McLeod gauge readings then fell from the steady value of 0·027 mm., obtained before heating the filament, to 0·019 mm., thus clearly indicating that steam was being formed as a result of a surface union of hydrogen and oxygen on the heated platinum wire. This was further confirmed by a progressive fall in the reservoir manometer readings as circulation proceeded. On interrupting the heating current, the wire immediately cooled and the McLeod gauge reading rose to its previous steady value of 0·027 mm. Evidently the rate of heat evolution during the formation of steam at this low pressure was not sufficient to compensate for the heat losses from the wire, which in consequence cooled and ceased to catalyse the reaction. Immediately the discharge was started, however, the platinum filament began to glow brightly, but cooled at once when the discharge was switched off. That the heating of the wire must have been due solely to the recombination of atomic hydrogen, and not to heating of the filament by the high-frequency field exciting the discharge, was shown by tha fact that the wire cooled immediately the discharge was interrupted by closing the throttle admitting gas into the discharge vessel.

The difference in behaviour of the platinum wire in atomic hydrogen and in a mixture of hydrogen and oxygen in equivalent proportions finds a natural explanation in the heat evolution of the appropriate reactions, the heat of formation of molecular hydrogen being 103,000 kcal/gm. mol, while that of steam is only about 59,000.

Further experiments were carried out with, in place of the platinum filament, a probe consisting of a straight length of either 36 S.W.G. tungsten or 37 S.W.G. copper wire protruding into the gas stream from a glass rod sealed into the port adjacent to the discharge vessel. When steam was fed to the discharge, these wires glowed dull red. With a mixture of steam and hydrogen in equal proportions, the copper wire became bright red and the tungsten wire was even hotter.

Wood (1922 a) also found that a tungsten wire glowed in atomic hydrogen ; he did not however obtain such an effect with a platinum wire. This may have been due to the state of the surface of his wire. It is a common experience that catalysing the combustion of a mixture of hydrogen and oxygen at a bright red heat not only cleans and activates a platinum surface but also roughens it.

Series 5—*Steam–mercury vapour mixtures.* At the conclusion of the experiments of Series 3 the question was left open as to whether or not the effect of the copper-staining of the discharge vessel walls in promoting a high yield of steam might be due, not only to the returning of molecular hydrogen to the reaction zone, but also to a surface reaction between atomic hydrogen and hydroxyl. In previous work on cathodic combustion (Finch 1935) it had been suspected that mercury vapour facilitated the combustion of hydrogen. This, and the fact that mercury vapour can play an important rôle in photosynthesized reactions involving hydrogen (Taylor and Bates 1926, Rieke 1936, Olsen 1938), suggested that its use might throw light on the present problem. Accordingly arrangements were made to examine the effect of mercury vapour additions to steam passing through the ring discharge.

About 0·5 cm³ of mercury was distilled into the discharge vessel port tube through a branch tube, which was then sealed off close to the port tube. A few turns of resistance wire wrapped round the port tube and immersed in an oil-filled jacket enabled the mercury to be warmed. An ash-coated probe was inserted through the first port beyond the discharge vessel. Mercuric oxide films were then formed, in the manner previously described, at the usual places where the circulating gas stream came in contact with stagnant mercury vapour. A similar deposit also formed about the entry to the discharge-vessel port tube.

Steam (or the equivalent hydrogen–oxygen mixture) was then circulated along the usual path under the normal pressure distribution of 0·026 mm. in the discharge vessel and 7 or 10 mm. in the reservoir, according to the nature of the gas mixture used. The mercury in the discharge port vessel was then gently heated and maintained at between 40 and 45° C. When steady conditions had been attained, the discharge was started. Instead of a rose-pink ring, it now formed an apparently uniform bluish-green sphere. A rough estimate afforded by the increased damping of the oscillation circuit suggested that the partial pressure of mercury vapour was between 0·005 and 0·012 mm., probably nearer the former. Despite this, no condensation occurred anywhere between the discharge vessel and the pump though, with time, a barely perceptible brownish band formed on the discharge vessel walls coaxial with the solenoid.

The mercury vapour exerted a profound effect on the results. The ash-coated probe did not incandesce and, although the mercuric oxide film near the discharge vessel port had disappeared within five minutes of starting the discharge, the more remote films were only very slowly reduced. Also, the equilibrium steam concentration, which previously had been between 18 and 19·4 per cent, was now practically 100 per cent, irrespective of the direction of approach to equilibrium. Finally, the spectrogram (Figure 4) of the discharge differs strikingly from that obtained with steam or steam–hydrogen mixtures (Figure 3). Apart from the strong, well-developed mercury line spectrum, its main features are the 2808 and 3064 A. hydroxyl bands, the latter being intense, and the three mercurous hydride line groups 3500, 4017 and 4219 A. Neither the secondary hydrogen nor the Balmer spectra appear.

These results show that the effect of the mercury vapour was largely to suppress the formation of atomic hydrogen, thereby shifting the equilibrium so far towards the steam side that the reaction products consisted almost wholly of steam. Thus the evidence suggests that steam is reduced, reversibly, by mercury to mercurous hydride and hydroxyl. In view of the high intensity of the 2536 A. mercury line [$Hg(6^3P_1)$] the reduction of mercury by hydrogen to mercurous hydride and atomic hydrogen is also likely to occur. Again, since mercury is oxidized by ozone to mercuric oxide, and this in turn is reduced in the cold by atomic hydrogen and by molecular hydrogen at about 120° c. (Manchot and Kampschulte 1907, 1910), further probable reactions are the oxidation of mercurous hydride to mercuric oxide and hydroxyl, and the reduction of mercuric oxide by atomic or molecular hydrogen to mercury and hydroxyl or steam respectively. In addition, the short life of hydroxyl is terminated in the manner set forth in Series 3 and 4 of the above experiments.

The question is still left open as to whether the survival life of hydroxyl is sufficient for it to interact with atomic hydrogen at the walls of the discharge vessel; the results of Series 4 show, however, that this reaction does not occur at surfaces beyond the discharge vessel.

Series 6—Dry and moist oxygen mixtures. The behaviour of oxygen, both dry and moist (or diluted with hydrogen), in various types of discharges, including the ring discharge, has been the subject of much experimental work. In general, a broad measure of agreement has been reached as to findings and conclusions. In one respect, however, a conflict of views does exist. Harteck and Kopsch (1931), Geib and Harteck (1934), Geib (1938) and others incline to the belief that, like atomic hydrogen, atomic oxygen has a relatively long life. On the other hand, Strutt (1910) and Finch and Bradford (1934) froze out ozone from the gas issuing from a discharge fed with dry oxygen, but found no evidence of the survival of atomic oxygen. All these workers appear to have used discharges between electrodes. Copeland (1930), however, used an electrodeless discharge, and concluded that atomic oxygen not only has a life comparable with that of atomic hydrogen but is also formed in the discharge by a mechanism similar to that postulated by Wood, whereby a poisoning of the discharge vessel walls by moisture was supposed to be necessary for the survival of the atomic species. It has, however, already been mentioned above that mercuric oxide films were formed after a brief circulation of oxygen through the discharge. These films were as readily obtained with dry as with moist oxygen. Thus if, as Copeland held, they were due to the interaction of mercury vapour and atomic oxygen, their formation with dry oxygen conflicted with his hypothesis. It was therefore decided to examine the matter further.

Owing to the difficulty of removing stains, caused by the experiments of Series 5, from the discharge vessel without etching the glass and thus impairing its smooth surface, a new vessel was fitted. It had previously been cleaned in the usual way but not baked out or degassed by circulating steam through the discharge. After fitting the first port beyond the discharge vessel with an ash-coated probe and the discharge vessel port with a 36 s.w.g. tungsten wire protruding into the discharge region, the normal circulation system was filled with oxygen previously washed with concentrated caustic potash and roughly dried by a single passage over phosphoric oxide. The gas was then circulated through the discharge under the normal conditions of pressure.

Rapid formation of the mercuric oxide films again occurred and was followed more slowly by a "tailing" of the mercury in the pump boiler. Neither the ash-coated nor the tungsten probe showed any signs of incandescence. The "tailing" and mercuric oxide films were easily removed by reduction, in the manner previously described, with atomic hydrogen formed from a hydrogen–steam mixture in equal proportions. A spectrogram, taken about half an hour after starting the discharge, is reproduced in Figure 5. It reveals, *inter alia*, a fairly strong atomic oxygen spectrum, faint hydroxyl bands, the first three lines of the Balmer series, and prominent bands due to carbon monoxide.

On repeating this experiment, a spectrogram similar to Figure 5 was obtained, but without the carbon monoxide spectrum. It would appear, therefore, that in the first experiment of this series the circulation of oxygen through the discharge had effectively out-gassed the discharge vessel walls. A fresh charge of oxygen was circulated for several hours past the heated spiral K and through the drying tube before starting the discharge, and a further spectrogram taken. This was again similar to Figure 5, except that, in addition to the absence of the carbon monoxide spectrum, the 3064, but not the 2808 A., hydroxyl band-head region was visible. H_β of the Balmer spectrum was very faint.

It will be evident at once that, in affording the primary (atomic) spectrum, the behaviour of dry oxygen in the ring discharge differed radically from that of dry hydrogen which, in the main, yielded the secondary spectrum, with weak hydroxyl bands and a few faint lines of the Balmer series attributable to traces of moisture.

It is noteworthy that, despite the spectrographic evidence of the production of atomic oxygen in the discharge, the tungsten and ash-coated probes remained cool throughout. Since the heat of formation of molecular oxygen ($>117,000$ kcal/gm. mol.) is greater than that of hydrogen, it is plain that little or no atomic oxygen survived to reach the ash probe, only 5 cm. beyond the discharge vessel, or even the tungsten wire probe. It is most unlikely, therefore, that the mercuric oxide films, which had formed so rapidly, could have been due to the interaction of mercury vapour and atomic oxygen, as Copeland has supposed. Rather must this phenomenon be ascribed to some other active modification of oxygen.

In the next experiments the effect of the dilution of oxygen by steam was examined. A wide range of mixtures was used containing between 5 and 96 per cent steam, the remainder being oxygen. The results have so much in common that it will suffice to summarize them as follows:

(i) Mercuric oxide films were formed, as with dry oxygen, at the usual sites, and "tailing" in the pump also occurred. The rate of formation of the mercuric oxide, as estimated from the growth in density of the film near the McLeod gauge, was roughly proportional to the oxygen concentration.

(ii) The ash-coated probe incandesced to some extent in all cases, but did so most vigorously in mixtures rich in steam. The tungsten wire probe glowed brightly in all cases.

(iii) The spectrograms reveal a progressive weakening of the atomic oxygen spectrum with increasing steam concentration, but a few lines, notably 3947 and 4368 A., remain, even in the case of pure steam. The hydroxyl band and Balmer spectra are prominent throughout.

It has already been shown in the case of dry oxygen that the formation of the mercuric oxide films was due to some active form of oxygen, other than atomic oxygen, which was incapable of exciting either the ash-coated or tungsten wire probe to incandescence. Hence, the fact that this oxidation process was slowed down by progressive dilution of the oxygen with steam strongly suggests that in the cases of both dry and moist oxygen the same agent was responsible for the oxidation of mercury. The results obtained by Strutt (1911) and Finch and Bradford (1934), taken into consideration with the fact that ozone is known to form yellow mercuric oxide with mercury (Manchot and Kampschulte 1907, 1910), justify the assumption that ozone was the long-lived, active form of oxygen produced in these experiments.

Since the ash probe and, to a lesser extent, the tungsten wire incandesced more brightly with increasing dilution with steam, it is evident that this could only have been due to recombination of atomic hydrogen. That the phenomena of the probe incandescence and mercuric oxide formation occurred simultaneously, even in mixtures containing as little as 5 per cent of oxygen, suggests that atomic hydrogen and ozone interact in the gas phase only relatively slowly, if at all, under the conditions of our experiments.

In the light of these results we are forced to conclude that, unlike atomic hydrogen, atomic oxygen does not survive as such on escaping from the discharge. There is a ready explanation for this. Thomson and Thomson (1928) point out that, unlike hydrogen, the oxygen molecule readily captures an electron to form a negative ion. It is to be expected that the oxygen atom is even more efficient in this respect. There is ample opportunity for such capture to occur within the discharge, where the electron population is of the same order as that of the molecules present (Thomson 1928). Under these circumstances, two-body collision reactions, such as the recombination of oxygen atoms to molecular oxygen, or of oxygen atoms with molecules to form ozone, are no longer subject to quantum-mechanical restrictions, provided one or both of the reactants is a negative ion.

Series 7—*Neon–steam mixtures.* This series of experiments forms, in a sense, a continuation of Series 6, the oxygen component being, however, replaced by a rare gas. It was hoped that a comparison of the results of the two sets of experiments would throw further light on, *inter alia*, the apparent strange reluctance of atomic hydrogen and ozone to interact. The experiments were carried out with mixtures covering a range from well-dried neon to pure steam.

The essential results can be summarized briefly by stating that (i) the intensity of the incandescence of the ash and tungsten wire probes increased with increasing steam concentration in a manner which closely paralleled that observed in the case of the corresponding oxygen–steam mixtures and, (ii) unlike the case of steam-diluted oxygen, the spectrograms (Figure 6) show that on increasing the steam concentration the neon spectrum was quickly suppressed. For instance, in a mixture of neon and steam in equal proportions only the hydroxyl and Balmer series spectra are visible.

The parallel behaviour of the probes in the two series supports the conclusion that in the case of both oxygen and neon–steam mixtures incandescence of the probes was caused solely by atomic hydrogen and, further, that atomic hydrogen and ozone are, for some unknown reason, slow to react with each other. It is also not difficult to see that, except for the formation or otherwise of mercuric oxide, the experimental results which led Copeland to suppose that atomic oxygen of long life is formed by the action of the electrodeless discharge on moist oxygen could equally well have been obtained with moist neon.

§4. SUMMARIZED FINDINGS AND CONCLUSIONS

The behaviour of steam or a mixture of hydrogen and oxygen in combining proportions, and their mixtures with hydrogen or oxygen, ranging from dry hydrogen to dry oxygen, have been studied in the ring discharge. The effect of mercury vapour and of neon on the steam equilibrium was also examined.

The principal facts established were as follows :

(i) *Dry hydrogen.* The secondary hydrogen spectrum was strongly emitted. The hydroxyl-band and Balmer spectra were weak and tended to disappear with progressive drying. Apart from traces attributable to incomplete drying, no agent capable of reducing yellow mercuric oxide at room temperature reached the walls of the discharge vessel.

(ii) *Moist hydrogen.* On adding to dry hydrogen about 1 per cent of oxygen or 2 per cent of steam (and possibly much less) the spectrum changed radically, the secondary hydrogen spectrum being replaced by the hydroxyl-band and Balmer spectra. Mercuric oxide at the wall of the discharge vessel was rapidly reduced.

(iii) *Steam.* The emission consisted mainly of the hydroxyl-band and Balmer spectra, together with a few lines due to atomic oxygen.

The stable end-products obtained from steam (or the equivalent hydrogen–oxygen mixture) consisted of steam, hydrogen and oxygen, the last two in their combining proportions. With a clean discharge vessel, the final steam concentration was between 16·5 and 19 per cent, but partial contamination of the vessel's walls by a barely visible film of copper sufficed to more than double the steam yield, which, however, returned to its previous value after removal of the copper film.

Glass needles, coated with sealing-wax ash (SnO and CaO) and inserted into the exhaust tube through ports situated 5 and 20 cm. beyond the discharge vessel, incandesced brightly. Films of yellow mercuric oxide deposited in the exhaust tube were rapidly reduced to metallic mercury. Copper, tungsten or platinum wire probes in the discharge vessel or in the exhaust tube glowed. The insertion of a copper wire spiral into the exhaust tube between the discharge vessel and an ash-coated probe suppressed the incandescence of the probe and practically stopped the reduction of the mercuric oxide films. The spiral did not appreciably affect the equilibrium steam concentration. With the ring discharge interrupted, an electrically heated platinum wire probe, situated in the exhaust tube near the discharge vessel, catalysed the formation of steam from the equivalent hydrogen–oxygen mixture, but ceased to do so and cooled when the heating current was stopped. The wire glowed spontaneously, however, immediately the ring discharge was restored.

(iv) *Steam–hydrogen and the equivalent hydrogen–oxygen mixtures.* The hydroxyl-band and Balmer spectra were prominent in all mixtures containing up to 95 per cent (and possibly more) diluent hydrogen, and their relative intensities appeared to be independent of such dilution. On increasing dilution with hydrogen, the atomic oxygen lines observed with steam alone weakened and eventually disappeared.

Dilution with hydrogen reduced the dissociation of steam, but not to the extent required by the law of mass action.

The ash-coated probes incandesced and the mercuric oxide films were reduced in all mixtures examined, incandescence being most pronounced in the case of mixtures containing hydrogen and steam in roughly equal proportions.

(v) *Steam–mercury vapour.* Apart from an intense mercury spectrum, the emission from the ring discharge in a mixture of steam (or the equivalent hydrogen–oxygen mixture) with roughly one-third its volume of mercury vapour consisted chiefly of the hydroxyl-band and mercurous-hydride spectra. The Balmer and secondary hydrogen spectra were absent. The ash probes did not incandesce.

The mercuric oxide films in the exhaust tube were only slowly reduced. In contrast to the steam yield of about 18 per cent obtained with pure steam, the equilibrium steam concentration rose to nearly 100 per cent.

(vi) *Dry oxygen.* The emission consisted mainly of the atomic oxygen spectrum, with faint hydroxyl-bands and a few faint lines of the Balmer series attributable to traces of moisture.

Yellow mercuric oxide was formed in the exhaust tube at all those sites where the gas stream issuing from the discharge came into contact with stagnant mercury vapour. The wire and ash-coated probes remained cool.

(vii) *Oxygen–steam.* The atomic oxygen spectrum weakened with increasing dilution by steam, but a few lines remained even with pure steam. The hydroxyl-band and Balmer spectra were prominent throughout.

Mercuric oxide films were formed, as with dry oxygen, even in mixtures containing only 5 per cent of oxygen. The ash-coated probes incandesced in all mixtures containing up to 95 per cent oxygen, but were brightest in steam-rich mixtures.

(viii) *Neon–steam.* The hydroxyl-band and Balmer spectra only gave way to the neon spectrum as the neon concentration approached 100 per cent. The behaviour of the ash-coated and wire probes was similar to that observed in the case of the corresponding oxygen–steam mixtures, in that the incandescence effect increased with increasing steam concentration.

The following conclusions have been drawn:

Dry molecular hydrogen is not dissociated but only excited in the ring discharge. The generally accepted view, according to which atomic hydrogen is formed and the absence of the Balmer spectrum is due to recombination of the atoms on the walls of the discharge vessel, is therefore untenable.

Atomic hydrogen issues in considerable quantity from the discharge in steam, hydrogen–steam or hydrogen–oxygen mixtures. The incandescence of ash-coated probes, and the spontaneous glowing of copper, tungsten, or platinum wires in the gases drawn from the discharge testify to a high concentration of atomic hydrogen, whilst the reduction of mercuric oxide films serves as a more sensitive test. It is unlikely that moisture adsorbed on the walls of the discharge vessel materially affects the yield of atomic hydrogen; for, if it did, it should also cloak the catalytic activity of a thin copper film, which it does not do.

The following reactions take place in hydrogen–steam and the equivalent hydrogen–oxygen mixtures. Steam is dissociated into hydroxyl and atomic hydrogen. The reverse reaction cannot proceed except through three-body collisions, and is therefore practically excluded. The same is true of the union of hydrogen atoms to molecules; consequently atomic hydrogen has a long survival life. On the other hand, the life of hydroxyl is short and is terminated by reacting with molecular hydrogen to form steam and atomic hydrogen, and with hydroxyl to form steam and atomic oxygen. Owing to the ease with which oxygen atoms form negative ions, they revert quickly to the molecular state. Finally, hydrogen reacts with oxygen to form steam and atomic oxygen. Some hydrogen peroxide may also be formed, but mainly at surfaces.

Dry oxygen, unlike dry hydrogen, is dissociated directly into atomic oxygen, which has a short life and quickly reacts within the discharge to form molecular oxygen and ozone. With stagnant mercury vapour, ozone forms films of yellow mercuric oxide. All the reactions occurring in oxygen and in hydrogen–steam mixtures also take place in oxygen–steam and the equivalent oxygen–hydrogen mixtures ; thus ozone and atomic hydrogen appear together in the effluent gases and interact, if at all, very slowly. The phenomena of the formation of mercuric oxide and the heating of ash-coated probes and wires in oxygen–steam mixtures are due to ozone and atomic hydrogen respectively, and are not evidence, as hitherto supposed, of a long life of oxygen in the atomic state.

The products leaving the ring discharge fed with a mixture of steam (or the equivalent hydrogen–oxygen mixture) and sufficient mercury vapour consist almost wholly of steam ; only traces of atomic hydrogen are present. In this case the steam is no longer directly dissociated into hydroxyl and atomic hydrogen, but reacts with mercury to form hydroxyl and mercurous hydride. Other reactions occurring are the oxidation of mercurous hydride to form mercuric oxide and hydroxyl, the reduction of mercuric oxide by atomic (or molecular) hydrogen to mercury and hydroxyl (or steam), the reduction of mercury by molecular hydrogen to mercurous hydride and atomic hydrogen, and the formation of steam and atomic oxygen from hydroxyl.

In neon–steam and the equivalent neon–hydrogen–oxygen mixtures, the reactions are similar to those occurring in steam.

ACKNOWLEDGMENTS

I have received much help in this study and wish to thank in particular Dr. E. E. Thompson and Dr. A. E. W. Austen, who carried out much of the experimental work, and Dr. R. W. B. Pearse, who checked the analyses of the spectrograms. My thanks are also due to Messrs. Ferranti for financial support of the work.

REFERENCES

Bonhoeffer, K. F., and Pearson, T. G., 1931, *Z. phys. Chem.* B**14**, 1.
Copeland, L. C., 1930, *Phys. Rev.*, **36**, 1221.
Esclangen, F., 1934, *Ann. Phys., Paris*, **11**, 267.
Finch, G. I., 1935, *J. Chem. Soc.*, 32.
Finch, G. I., and Bradford, B. W., 1934, *J. Chem. Soc.*, 360.
Finch, G. I., and Fraser, R. P., 1926, *J. Chem. Soc.*, 117.
Finch, G. I., and Mahler, E. A. J., 1931, *Proc. Roy. Soc.* A, **133**, 173.
Frost, A. A., and Oldenberg, O., 1936, *J. Chem. Phys.*, **4**, 642.
Geib, K. H., 1938, *Z. Elektrochem.*, **44**, 81.
Geib, K. H., and Harteck, P., 1934, *Trans. Faraday Soc.*, **30**, 134.
Harteck, P., and Kopsch, U., 1931, *Z. phys. Chem.*, B**12**, 327.
Mackinnon, K. A., 1929, *Phil. Mag.*, **3**, 605.
Manchot, W., and Kampschulte, W., 1907, *Chem. Ber.*, **40**, 2891 ; 1910, *Ibid.*, **43**, 750.
Mott, N. F., and Massey, H. S. W., 1933, *Atomic Theory of Collisions* (Oxford : Clarendon Press), p. 214.
Oldenberg, O., 1937, *J. Phys. Chem.*, **41**, 293.
Olsen, L. O., 1938, *J. Chem. Phys.*, **6**, 307.
Poole, H. G., 1937, *Proc. Roy. Soc.* A, **163**, 404.
Rieke, F. F., 1936, *J. Chem. Phys.*, **4**, 513.
Southworth, G. C., 1920, *Radio Rev.*, **1**, 583.
Steacie, E. W. R., 1946, *Atomic and Free Radical Reactions* (New York : Reinhold).
Strutt, R. J., 1910, *Proc. Phys. Soc.*, **23**, 66 ; 1911, *Proc. Roy. Soc.* A, **85**, 226.
Taylor, H. S., and Bates, J. R., 1926, *Proc. Nat. Acad. Sci., Wash.*, **12**, 714.
Thomson, J. J., 1928, *Proc. Phys. Soc.*, **40**, 79.
Thomson, J. J., and Thomson, G. P., 1928, *Conduction of Electricity through Gases* (Cambridge : University Press), 3rd ed., p. 242.
Wood, R. W., 1920, *Proc. Roy. Soc.* A, **97**, 455 ; 1921, *Phil. Mag.*, **42**, 729 ; 1922 a, *Ibid.*, **44**, 538 ; 1922 b, *Proc. Roy. Soc.* A, **102**, 1.

The Significance of the Observation of Intense Radio-Frequency Emission from the Sun

By M. RYLE
Cavendish Laboratory, Cambridge

Communicated by Sir Lawrence Bragg ; MS. received 7th April 1949

ABSTRACT. A summarizing account is given of possible mechanisms which have been proposed to account for the emission of intense radio waves from the sun. Whilst most authors agree that the radiation from the undisturbed sun is due to an electron temperature of a million degrees in the solar corona, the greater intensity emitted from sunspots has led some authors to propose non-thermal mechanisms, based on the coherent oscillation of a large number of electrons in the solar corona above the sunspot. The mechanisms proposed for the maintenance of coherent oscillations are examined in detail in this paper and it is concluded that whilst they give adequate explanations for the oscillation observed in discharge tubes, no satisfactory mechanism has yet been suggested for the maintenance of electron oscillations in the solar corona.

It is therefore concluded that the observations can only be accounted for by the occurrence, in the solar corona near sunspots, of electron temperatures of up to 10^{10} deg. K.

The problem of maintaining an electron temperature of 10^6 deg. K. in the normal corona, and 10^{10} deg. K. in the corona near a large sunspot, has been considered in an earlier paper ; it was there shown that by the action of the general magnetic field of the sun and the non-uniform rotation of the solar surface, a potential difference is developed between the poles and the equator which is capable of maintaining an adequate electron temperature in both regions.

§1. INTRODUCTION

RECENT experimental work has shown that the sun emits radio waves of great intensity. For wavelengths between 1 metre and 5 metres, the intensity corresponds to black-body emission at temperatures of about 10^6 deg. K. for the undisturbed sun and up to 10^{10} deg. K. for regions near sunspots. There is general agreement that the radiation from the undisturbed sun is due to a genuine electron temperature of about a million degrees in the solar corona.

Some authors have suggested that the more intense radiation from regions near sunspots is produced by the coherent oscillation of a large number of electrons by some mechanism similar to that which produces plasma oscillations. If this were the true explanation it would not be necessary to assume that, near sunspots, the electrons had a mean energy corresponding to a temperature of 10^{10} deg. K.

The main object of this paper is to examine in detail the mechanisms which have been proposed for the coherent oscillation of electrons in the solar corona. It is shown that although it is a simple matter to demonstrate the occurrence in the solar corona of a natural " plasma " frequency of the electron gas, it is much more difficult to suggest any mechanism which would maintain the plasma oscillations, or any mechanism by which radio waves could be emitted from the plasma even if it did oscillate. All the mechanisms proposed to explain the oscillations in discharge tubes require sharp gradients of electron density and electron beams having a highly ordered motion ; it is difficult to see how either of these conditions could occur in the solar corona. The conclusion is therefore reached that no satisfactory mechanism has yet been suggested. It is therefore concluded that the radiation emitted from near sunspots must be produced by the non-coherent acceleration of electrons, so that its observation implies the occurrence of genuine electron temperatures of 10^{10} deg. K. in these regions. A mechanism for the maintenance of such a great electron temperature in the neighbourhood of sunspots has been considered in an earlier paper. It was there suggested that by the action of the general magnetic field of the sun and the non-uniform rotation of the solar surface, a large potential difference is developed between the poles and the equator, which is capable of maintaining a mean electron energy in the undisturbed corona corresponding to a temperature of at least 10^6 deg. K. The distortion of the general magnetic field by the magnetic field of a sunspot allows the maintenance of still greater electric fields in the corona near the sunspot, and mean electron energies corresponding to a temperature of 10^{10} deg. K. are likely to occur near large sunspots.

§ 2. THE EXPERIMENTAL OBSERVATIONS

The earliest deductions that the sun occasionally emits radio waves whose intensity greatly exceeds that expected from a source at 6,000° K. were made by Appleton (1945) and Hey (1946) from observations on frequencies of about 70 Mc/s. Since that time a large amount of experimental work has been carried out on frequencies between 60 Mc/s. and 3,000 Mc/s. It has been found that for frequencies between 60 Mc/s. and 200 Mc/s. the intensity from the undisturbed sun corresponds to what would be emitted if the whole solar disc were at a temperature of about a million degrees (Pawsey 1946, Ryle and Vonberg 1948).

At times of sunspot activity much greater intensities have been observed (Appleton 1945, Hey 1946, Pawsey, Payne-Scott and McCready 1946, Ryle and Vonberg 1946); the radiation at frequencies of about 200 Mc/s. then corresponds to what would be emitted if the whole solar disc were at a temperature of about 10^8 deg. K.

Experiments have also been made to determine the position and size of the source of the intense radiation (Ryle and Vonberg 1946, McCready, Pawsey and Payne-Scott 1947). The results obtained have shown that the radiation is emitted from a comparatively small region in the neighbourhood of the sunspot. The observed intensities of radiation therefore correspond to a still greater effective temperature, and it must be concluded that regions near sunspots are capable of emitting radiation whose intensity corresponds to the emission from a black-body source at a temperature exceeding 10^{10} deg. K.

The intense radiation from sunspots has been found to be circularly polarized, but the radiation from the undisturbed sun is apparently unpolarized (Martyn 1946 a, Appleton and Hey 1946, Ryle and Vonberg 1946, 1948).

§ 3. EQUILIBRIUM MECHANISMS LEADING TO THE DEDUCTION OF VERY HIGH TEMPERATURES

Some theories of radio-frequency emission from the sun have been based upon "equilibrium mechanisms", in which the intensity of the radiation tends to the value associated with the mean electron energy. The radiation from an infinite heated ionized gas of low density would correspond to that from a black-body, although in practical cases the total "optical depth" may be insufficient for the full intensity to be established. It is clear that such theories cannot account for radiation of intensity greater than that emitted by a black-body having the same temperature as that of the electrons; an explanation of solar radiation by these theories therefore requires the existence of very high electron temperatures.

(i) *Undisturbed Sun*

In a previous communication (Ryle 1948) a mechanism for the generation of radio-frequency radiation in the solar atmosphere was developed. By an application of Appleton's magneto-ionic theory to the propagation of metre waves in the solar atmosphere, it was shown that radiation could only escape to the earth from a source situated at or above the level at which the refractive index vanished (about 5×10^9 cm. above the photosphere for a wave of frequency 200 Mc/s.). It was further shown that the absorption of an incoming wave was negligible except near the level at which the refractive index vanished, but that the total absorption in this region was considerable. It was therefore concluded that each region in the solar atmosphere would radiate at a particular frequency with an intensity which might approach that corresponding to a black-body radiator having the same temperature as the electrons in the region. It was therefore suggested that the observed intensity of the radiation from the undisturbed sun was the result of an electron temperature of at least a million degrees in the corona.

It was suggested that the maintenance of such an electron temperature could be accounted for by the existence of a potential difference between the poles and equator of the sun, the potential difference being developed by the action of the general solar magnetic field on the non-uniform rotation of the photosphere.

Martyn (1946 b, 1948) has also developed a "thermal" theory of the radiation from the undisturbed sun, and whilst not suggesting a mechanism for the maintenance of the electron temperature, he considers that the observations can be accounted for by assuming a temperature of a million degrees in the corona. Waldmeier (1948) and Shklovsky (1947) have also developed theories involving a comparable coronal temperature. Moreover, the observation of lines of very highly ionized atoms in the coronal spectrum leaves little doubt that the radiation from the undisturbed sun is due to purely "thermal" processes in the corona.

(ii) *Sunspot Radiation*

In the previous paper (Ryle 1948) the theory was extended to the case of sunspot radiation. It was shown that in the solar corona above a sunspot appreciable absorption could occur at two levels (about 10^9 cm. and 10^{10} cm. above the photosphere for a frequency of 200 Mc/s. and a sunspot magnetic field of 3,000 gauss); each region would absorb one of the circularly polarized components. It was therefore suggested that circularly polarized radiation would be generated in each of these regions; the intensities of the two waves would approach the values associated with the electron temperatures in the two regions.

It was further shown that the distortion of the general solar magnetic field by the magnetic field of the sunspot could give rise to a very much greater electric field in the corona above a sunspot than in the undisturbed corona. It was shown that if the non-uniform rotation of the sun produced a potential difference between poles and equator which was capable of maintaining a normal electron temperature of 10^6 deg. K., then the distortion of the magnetic field near a large sunspot might be expected to lead to temperatures of the order of 10^{10} deg. K.

§4. MECHANISMS INVOLVING THE COHERENT OSCILLATION OF ELECTRONS

(i) *Non-thermal Theories of Sunspot Radiation*

The extremely great electron temperatures which are required to account for the observed intensity of the radiation from sunspots have led other authors to propose non-thermal mechanisms. In order to produce an intense radiated field by electrons of comparatively low energy, it seems necessary to develop a mechanism for maintaining the coherent oscillation of a large number of electrons.

Martyn (1947) and Shklovsky (1947) have both suggested that the coherent oscillation of electrons in the corona could be produced by a mechanism analogous to the plasma oscillations familiar in gas discharge tubes. If such a process were to occur, the intensity of the radiation emitted could be very much greater than that associated with the thermal motion of the electrons in the corona. It is therefore important to examine whether oscillations of this kind could occur under the conditions likely to exist in the solar corona.

(ii) *Existing Theories of the Mechanism of Plasma Oscillations*

Before attempting to consider the possibility that plasma oscillations could account for the intense emission from the solar corona it is important to understand the precise mechanisms which are envisaged in such theories. Unfortunately, the literature of plasma oscillations is notable for the small amount of detailed theoretical study as compared with the quantity of careful experimental results which have been obtained in the study of gas discharges.

In the original theory of Tonks and Langmuir (1929) it was shown that the electrons in an ionized gas had a natural frequency of oscillation given by $(N\mathbf{e}^2/\pi\mathbf{m})^{\frac{1}{2}}$, where N is the electron density and $\mathbf{e}$ and $\mathbf{m}$ the charge and mass of the electron. (They also showed the existence of ionic oscillations having much lower frequencies, but it will be seen later that such oscillations cannot account for the emission of radio-frequency radiation from the sun.) It should be noted that although this theory provided natural periods of oscillation of the same order as those observed in experiments on gas discharges, no definite mechanism was proposed for the excitation of such oscillations. Without some additional mechanism, a plasma is not capable of generating oscillations, and has merely the characteristics of a tuned circuit. If some feedback mechanism or "negative resistance" is introduced, however, a plasma will then become capable of generating oscillations.

A further difficulty arises in accounting for the detection of the oscillations. This difficulty is due to the fact that the group velocity in the plasma is zero (Tonks and Langmuir 1929, Tonks 1929); it is therefore impossible for energy to be propagated by an electromagnetic field from the oscillating region to the measuring probe, unless the gradient of electron density or magnetic field is sufficiently great to allow appreciable transfer by way of the evanescent wave.

The original theory has since been extended to account both for the excitation of plasma oscillations and for the escape of energy from the oscillating region.

(*a*) *The excitation of plasma oscillations.*

Later work by Merrill and Webb (1939) suggested that the oscillation of a thin layer near the cathode was responsible for velocity modulation of the primary electron stream from the cathode, a conclusion which has been confirmed by

Armstrong (1947). Recently Neill (1949) has shown that it is possible to account for the experimental observations by a feedback mechanism analogous to that by which a diode can maintain oscillations in virtue of transit-time effects. He supposes that electrons emitted by the cathode are subject to an accelerating field which depends on the space-charge in the plasma-sheath; if this space-charge is varying, the electron stream will be velocity-modulated, and if the distance to the plasma is suitable it will produce a bunching of the electrons which in turn can affect the space-charge. In this way part of the energy from the cathode fall of potential is converted into energy of oscillation of the plasma. Neill has shown that the theory provides good quantitative agreement with many of the related observed phenomena.

A more elaborate mechanism for the generation of oscillations in vacuum tubes has recently been suggested by Haeff (1949), which depends upon the interaction of two coherent beams of electrons.* Haeff has shown that under certain conditions a region containing two interpenetrating electron beams is capable of amplifying an initial disturbance having a frequency related to the plasma frequencies of the two beams. Whilst insufficient detail has been given of the precise mechanism, it appears that it is analogous to that of the travelling-wave tube, in that the disturbance is guided along the moving electron beams, and that the wave frequency to an observer moving with the electron beam is related to the plasma frequency.

This brief survey shows that all the detailed mechanisms so far proposed for the maintenance of oscillations in a gas discharge tube depend on the velocity modulation of an incident electron beam. It will be necessary to examine (in § 4 (iii)) whether such a mechanism could provide excitation for the coherent oscillation of electrons in the solar corona.

(*b*) *The escape of energy from the oscillating region.*

Tonks and Langmuir (1929) considered that even in the small-scale conditions of the discharge tube, where the entire plasma occupies a thickness considerably less than a vacuum wavelength, it might be necessary to account for the detection of the oscillations by virtue of the escape of accelerated electrons from the plasma region. Merrill and Webb (1939) concluded that the alternating voltage observed on the measuring probe in their experiments was due to the periodic fluctuations of the space-charge in the incident electron stream. Later experiments by Armstrong (1947) and Neill (1949) have shown that the oscillating plasma will cause velocity modulation of the incident electron stream; they have also confirmed the earlier conclusion that the voltage observed on the measuring probe is caused by the bunching of such an incident electron stream.

Velocity modulation of the incident stream of electrons could also account for the escape of energy from a region which was oscillating by the mechanism suggested by Haeff (1949). The possibility that electromagnetic radiation could escape from such an oscillating region cannot be excluded until a more precise physical picture of the mechanism has been given; since, however, the emission of an electromagnetic wave relative to a coordinate system moving with the stream of electrons corresponds closely to the escape of radiation from a normal plasma, it seems probable that the escape of energy by electromagnetic radiation presents considerable difficulty.

---

* A mechanism somewhat analogous to that proposed by Pierce (1948) to account for low-frequency oscillations in an electron stream passing through a stationary ion cloud.

In the absence of further information it therefore appears that energy is only likely to escape from an oscillating plasma by virtue of the velocity modulation imposed on an electron stream which has passed through the region.

It is clear that in the large-scale conditions of the solar corona, the escape of energy from an oscillating plasma presents even more serious difficulties than for the case of the discharge tube; the application of the theories outlined above to the solar corona will be discussed in the next section.

(iii) *The Possibility of Electronic Plasma Oscillations in the Solar Corona*

It is first important to note that the existing experimental and theoretical work on plasma oscillations in gas discharge tubes relates to conditions in which the dimensions of the plasma are small compared with the vacuum wavelength of the oscillation. In the solar corona it is probable that the electron density and magnetic field will be virtually constant over a volume of many cubic wavelengths.

(*a*) *The excitation of plasma oscillations in the corona.*

In a gas discharge tube containing a cathode at a low temperature, the initial velocities of the electrons are small compared with the velocity produced by the applied electric field. Under the influence of a steady applied field, the stream of electrons will therefore travel a distance of the order of one mean free path before the motion becomes disordered. If now an alternating field is superimposed on the steady field near the cathode, the electrons will be velocity-modulated and, if the mean free path is sufficient, will become bunched at some distance from the cathode. Neill (1949) has shown that the variation of the space-charge resulting from this bunching can then provide excitation to the oscillating plasma.

If now the temperature of the cathode is increased, the initial distribution of the velocities of the electrons may become comparable with the velocity produced by the applied field; although the superimposition of an alternating field will still result in the acceleration of each electron, the spread in the initial velocities of the electrons will prevent effective bunching. An alternating density of space-charge will therefore no longer be produced, and the oscillations of the plasma cannot be maintained.

In the solar corona the initial random velocities of the electrons in any particular region will be of the same order as the velocities gained during a mean free path, and the conditions therefore correspond to those in a discharge tube having a cathode at a very high temperature. The superimposition of an alternating field will not produce appreciable bunching in such circumstances; the mechanism suggested by Neill (1949) cannot therefore account for the maintenance of plasma oscillations in the solar corona.

Although Haeff (1949) has given no detailed description of the conditions which are necessary for the generation of oscillations by the interaction of two electron beams, it seems likely that here also the distribution of velocities in the primary electron beams must initially be small compared with the ordered velocity.

It appears that any mechanism which involves velocity modulation of an incident electron stream is equally unlikely to be effective in maintaining the coherent oscillation of electrons in the conditions of the solar corona.

(*b*) *The escape of the energy.*

Even if it is assumed that the coherent oscillation of electrons could be maintained in the solar corona, it is doubtful whether any energy so generated could escape to the earth. It is clearly more difficult to account for leakage of the evanescent wave in a region where the gradient of electron density and magnetic field will be very much smaller than those in a discharge tube. It is therefore necessary to examine the possibility that energy might be transferred from the oscillating region by the passage of electrons through the region. A stream of electrons undergoing velocity modulation in this way might pass into an overlying region from which radiation could escape; such a stream is unlikely to be an efficient radiator, but some of the energy imposed by the oscillating plasma might be able to escape. It is, however, necessary that the electron stream shall still be effectively bunched when it emerges into the region from which the radiation can be propagated. Unless there are very considerable gradients of electron density or magnetic field, the electron stream must travel a distance corresponding to a very large number of wavelengths. The accurate collimation of a velocity-modulated electron stream which has traversed such a distance can only occur if the random velocity of the electrons is negligible compared with their ordered motion. It has already been shown that such a condition is unlikely in the solar corona.

It therefore appears that although the electrons in the solar corona may have a suitable natural period of oscillation, there are serious difficulties in accounting for the intense radiation at radio frequencies by mechanisms involving plasma oscillations. Satisfactory theories have been developed to account for both the driving mechanism and the mechanism of energy transfer for plasma oscillations in discharge tubes; these theories are not applicable in the large-scale conditions of the solar corona unless very large gradients of electron density or magnetic field are assumed. The maintenance of such gradients in a region where the mean free path for an electron is many kilometres does not appear possible.

(iv) *The Possibility of Ionic Plasma Oscillations in the Solar Corona*

So far attention has only been given to the coherent oscillation of electrons. If positive ions are also present, they will be capable of oscillating with a natural frequency of $(N\mathbf{e}^2/\pi M)^{\frac{1}{2}}$, where N and M are the number density and mass of the ion and $\mathbf{e}$ is the electronic charge (Tonks and Langmuir 1929). In a region containing an equal number of ions and electrons, the natural frequency of the ionic oscillations will be very much less than that of the electronic oscillations; the refractive index for a wave having the frequency of the ionic oscillations is therefore imaginary, and ionic oscillations cannot generate radio-frequency radiation which can escape through the overlying layers.

Tonks and Langmuir (1929) have, however, also shown that oscillations of very low frequency can be generated in an ionized gas, and that these oscillations may be propagated in a manner analogous to the propagation of sound waves. It is possible that such oscillations might occur in the solar corona; the passage of such a wave through a region which was emitting radio-frequency radiation would produce variations in the mean energy or density of the particles, and might cause fluctuations in the intensity of the emission. Thus, whilst ionic plasma oscillations cannot account for the great intensity of the radio-frequency radiation emitted by the corona, it is possible that they might produce low-frequency fluctuations in the intensity of the emitted radiation. Such waves might be responsible for the "bursts" of radiation associated with the emission from sunspot regions.

§ 5. CONCLUSIONS

A detailed consideration of theories of plasma oscillations has not revealed a satisfactory explanation for the maintenance of coherent electronic oscillations in the solar corona. It is therefore concluded that the observed radio-frequency radiation arises from purely random motion of the electrons in the corona. It has previously been shown (Ryle 1948) that each region of the corona is absorbing to a certain radio frequency, and thus each region will radiate at a particular frequency with an intensity depending upon the electron temperature and the total absorption coefficient. If the absorption is sufficient, the intensity will equal that of black-body radiation corresponding to the electron temperature in the region.

If these conclusions are accepted, the observations show that the electron temperature in the normal corona is at least a million degrees (a figure assumed in many theories of the radiation from the undisturbed sun), whilst in the corona above large sunspots the electron temperature may reach 10^{10} deg. K. The existence of such electron temperatures in the solar corona is clearly of great significance; observations over a range of radio frequencies may allow a detailed study to be made of the distribution of electron temperature throughout the corona.

A mechanism for the maintenance of a large electron temperature in the solar corona has already been suggested (Ryle 1948). It was shown that the non-uniform rotation of the solar surface, together with the general magnetic field of the sun, will produce a potential difference of about 5×10^8 volts between poles and equator. A potential difference of this order is sufficient to maintain not only a normal electron temperature in the corona of about a million degrees, but also an increased electron temperature in the corona above sunspots. For a large sunspot the potential gradient is sufficient to maintain in the regions responsible for the emission of radio-frequency radiation a mean electron energy corresponding to a temperature of 10^{10} deg. K.

The application of similar mechanisms to account for the discrete sources of radio waves in the galaxy, and the possible relation between these sources and the origin of cosmic rays, will be considered in another paper (Ryle 1949).

ACKNOWLEDGMENTS

I am indebted to Mr. J. A. Ratcliffe for his advice and helpful criticism throughout this work. I also had the advantage of useful discussions with Professor K. G. Emeléus and Professor J. Sayers on the mechanism of plasma oscillations.

REFERENCES

Appleton, E. V., 1945, *Nature, Lond.*, **156,** 534.

Appleton, E. V., and Hey, J. S., 1946, *Nature, Lond.*, **158,** 339.

Armstrong, E. B., 1947, *Nature, Lond.*, **160,** 713.

Haeff, A. V., 1949, *Proc. Inst. Radio Engrs.*, **37,** 4.

Hey, J. S., 1946, *Nature, Lond.*, **157,** 47.

McCready, L. L., Pawsey, J. L., and Payne-Scott, R., 1947, *Proc. Roy. Soc.* A, **190,** 357.

Martyn, D. F., 1946 a, *Nature, Lond.*, **158,** 308 ; 1946 b, *Ibid.*, **158,** 632 ; 1947, *Ibid.*, **159,** 26 ; 1948, *Proc. Roy. Soc.* A, **193,** 44.

Merrill, H. J., and Webb, H. W., 1939, *Phys. Rev.*, **55,** 1191.

Neill, T. R., 1949, *Nature, Lond.*, **163**, 59.
Pawsey, J. L., 1946, *Nature, Lond.*, **158**, 633.
Pawsey, J. L., Payne-Scott, R., and McCready, L. L., 1946, *Nature, Lond.*, **157**, 158.
Pierce, J. R., 1948, *J. Appl. Phys.*, **19**, 231.
Ryle, M., 1948, *Proc. Roy. Soc.* A, **195**, 82 ; 1949, *Proc. Phys. Soc.* A, **62**, 491.
Ryle, M., and Vonberg, D. D., 1946, *Nature, Lond.*, **158**, 339 ; 1948, *Proc. Roy. Soc.* A, **193**, 98.
Shklovsky, J. S., 1947, *Nature, Lond.*, **159**, 752.
Tonks, L., 1929, *Phys. Rev.*, **33**, 239.
Tonks, L., and Langmuir, I., 1929, *Phys. Rev.*, **33**, 195.
Waldmeier, M., 1948, *Experientia*, **4**, 64.

Evidence for the Stellar Origin of Cosmic Rays

By M. RYLE
Cavendish Laboratory, Cambridge

Communicated by Sir Lawrence Bragg ; MS. received 7th April 1949

ABSTRACT. Several authors have suggested the possibility that cosmic rays are due to the acceleration of charged particles in the atmospheres of certain stars. Recent observations at radio frequencies have provided some evidence in support of this hypothesis which is reviewed in this paper.

The observations have shown that at least part of the radio-frequency radiation from the galaxy is emitted by discrete sources of small angular diameter. Measurements of the intensity of the radiation from these sources, and of the time variations of the intensity, indicate that the sources emit radio waves as if they were at a temperature of more than 10^{14} deg. K. It is concluded that this intense radiation cannot be caused by the coherent oscillation of a large number of electrons, but must be due to a genuine electron temperature of about 10^{14} deg. K. (corresponding to a mean electron energy of 10^{10} electron volts). The existence of this electron temperature indicates that in certain stellar bodies there are mechanisms capable of accelerating particles to cosmic ray energies.

A previous theory of the emission of radio waves from the sun is extended to other stellar bodies. It is concluded that a star having a surface magnetic field strength and a peripheral velocity somewhat greater than those which have so far been observed could emit radio waves having the observed intensity. A star of this type could also accelerate charged particles to cosmic ray energies.

The small visual brightness of the sources of radio waves suggests that these bodies may be characterized by small visual opacity or small photospheric temperature.

It is suggested that the difference between the angular distribution of cosmic rays and of the radio waves from the galaxy is due to the deflection of cosmic ray particles by magnetic fields. In addition to previous theories of a general magnetic field in the galaxy, it is shown that the scattering produced by the magnetic fields of the distributed sources themselves should be sufficient to produce an isotropic distribution of cosmic rays at the earth even though the sources may show a marked concentration near the galactic equator.

§1. INTRODUCTION

Several authors have proposed mechanisms for the generation of cosmic rays by the acceleration of charged particles in the atmospheres of certain stars. Swann (1933) demonstrated that the growth of the magnetic field of a sunspot could give rise to potentials of the order of 10^9 volts in a region where little energy would be lost by collisions. He suggested that similar disturbances in other stars might give rise to particles having an energy of more than 10^{10} electron volts. Alfvén (1937) showed that even in the absence of sunspots a comparable potential could be developed by the non-uniform rotation of the solar surface. Davis (1947) deduced that the star BD-18°-3789, which recent measurements (Babcock 1947b) have shown to have a magnetic field strength at the poles of about 5,500 gauss, would produce, relative to a stationary observer, a potential difference of about 10^{14} volts. He showed, however, that owing to the formation of a space-charge this potential difference was not available for the acceleration of charged particles. Babcock (1948) has considered this process in relation to magnetically variable stars, and he has shown that the periodic reversal of the magnetic field would allow the escape of particles of high energy during a certain fraction of the magnetic cycle.

Recent measurements (Hey, Parsons and Phillips 1946, Bolton and Stanley 1948, Ryle and Smith 1948) on the emission of radio-frequency radiation from the galaxy have provided independent evidence for the existence of such sources; it is the purpose of this paper to examine this evidence and to make certain deductions concerning the nature of the sources.

The observations have shown that at least part of the radio-frequency radiation from the galaxy is emitted by discrete sources, whose angular diameters are smaller than the resolving power of the apparatus (about eight minutes of arc). It has also been found that the intensity from some of the sources fluctuates rapidly and that sometimes "bursts" of radiation are observed, in which the intensity is doubled for a period of about 20 seconds. Providing that these fluctuations represent genuine variations of the emission from the source, it is possible to fix an upper limit to the dimensions of the source. Extended measurements of the positions of the two intense sources have suggested that they are not part of the solar system; if it is assumed that they are situated at the distance of the nearest fixed star, the deduced diameter and the observed intensity would correspond to the emission from a black-body source at a temperature of more than 10^{14} deg. K.

It is possible that some special radiating process might be responsible for the emission of radiation having an intensity greater than that corresponding to the electron temperature; in the absence of such a process, the radiation can only be accounted for by assuming that the electrons have a random energy corresponding to the thermal energy at a temperature of 10^{14} deg. K. (the mean energy of an electron is then about 10^{10} electron volts).

The maintenance of this electron energy implies the existence of mechanisms capable of accelerating particles to cosmic ray energies. It is therefore suggested that the sources of intense radio-frequency radiation may also be responsible for the emission of cosmic rays.

The assumptions made in the development of this theory are analysed in detail in the following sections. A mechanism is suggested by which intense radio waves (and cosmic ray particles) could be emitted from the atmospheres of certain stars having a large rotational velocity and surface magnetic field. It is further shown that the scattering effect caused by the magnetic fields of such stars would be expected to lead to an isotropic distribution of cosmic rays at the earth, independent of the distribution of the sources in the galaxy.

§2. THE EXPERIMENTAL EVIDENCE

Hey, Parsons and Phillips (1946) first observed that the intensity of the radio-frequency radiation from the constellation of Cygnus showed fluctuations in a period of a few minutes; they therefore concluded that at least part of the radiation from the galaxy was due to the emission from discrete sources. More refined experiments by Bolton and Stanley (1948) and Bolton (1948) enabled the source to be located more precisely, and showed that its angular diameter was less than the resolving power of the apparatus (about eight minutes of arc). Similar observations by Ryle and Smith (1948) showed a second, more intense, source in the constellation of Cassiopeia, and extended observations showed that whilst the intensity from the source in Cygnus frequently showed sudden increases, that from the source in Cassiopeia was notably constant. They also found that the radiation from both sources appeared to be randomly polarized.

The intensity of the 80 Mc/s. radiation from the sources in Cygnus and Cassiopeia has been found to be of the order of 2×10^{-22} w.m^{-2}(c/s.)$^{-1}$. Both Bolton (1948) and Ryle and Smith (1948) have detected other discrete sources of smaller intensity, and some of these have also shown fluctuations of intensity. None of the sources appears to coincide with an outstanding visual star.

§ 3. DETERMINATION OF THE EQUIVALENT TEMPERATURE OF A SOURCE

The intensity of the radiation emitted by a distant source may be compared with that from a black-body source which subtends an equal solid angle. If the intensity corresponds to that produced by the black-body at a temperature T, the source may be said to have an "equivalent temperature" T. The significance of this temperature will be discussed in § 4, and it will be used now merely as a convenient unit with which to express the emission from the source.

In order to determine the equivalent temperature of a galactic source of radio-frequency radiation, it is necessary to observe both the intensity of the radiation at the earth and the solid angle subtended by the source. Unfortunately it has not been possible to achieve a sufficiently large resolving power to measure the solid angle directly. Observations of the source in Cygnus have, however, shown sudden fluctuations in which the intensity is doubled in a period of 20 seconds. This variation might represent variations in emission, or might be caused by refraction in an intervening medium; it will first be shown that a refraction process is unlikely.

The observation that the intensity from the source in Cygnus exhibits rapid fluctuations, whilst that from the source in Cassiopeia remains noticeably constant, suggests that the fluctuations are not produced by general refraction in interstellar matter. Since the observations were made when the sources were nearly vertically overhead, it seems unlikely that refraction in the ionosphere could affect one source and not the other (the angles of incidence during observation are about 12° for the source in Cygnus and 6° for the source in Cassiopeia). Whilst the possibility still remains that the intensity from the source in Cygnus is affected by refraction in some localized region of interstellar matter, it seems most probable that the observed fluctuations represent a true variation of the emission.

The total intensity of the radiation received from a source may be considered as being due to the addition of the radiation emitted by a large number of small elements of the source. If the radiation from each of these small elements exhibits sudden increases of intensity, the total intensity will only show such increases if the fluctuations from different elements arrive simultaneously. If large increases are observed it is therefore necessary that the different elements of the source shall be separated by a distance not greater than the distance travelled by an electromagnetic wave in the duration of the fluctuations. In the absence of any special orientation of the emitting surfaces relative to the earth it is therefore possible to deduce a maximum dimension of the source by determining the duration of the sudden fluctuations.

Observations of the source in Cygnus have provided records showing a doubling of the intensity in a time of 20 seconds. It is therefore deduced that the source of the variable component of the radiation occupies a volume whose linear dimensions do not greatly exceed the distance travelled by electromagnetic waves in 20 seconds (about 10^{12} cm.).

There is, as yet, insufficient information on the distance of the sources. Parallax measurements at radio frequencies are unlikely to be of sufficient accuracy to determine the distances, although preliminary measurements by Ryle and Smith (unpublished) have shown that the sources in Cygnus and Cassiopeia are at distances greater than 2×10^{16} cm. Improvements in the angular accuracy of the measuring apparatus are unlikely to lead to a significant extension of parallax measurements by radio methods, although they may allow the sources to be identified with known stellar bodies, in which case parallax measurements by visual methods may be possible. The existing measurements indicate that it is unlikely that the sources are part of the solar system, and in the absence of further experimental evidence it will be assumed that they are not nearer than the nearest known fixed star (about 3×10^{18} cm.).

With this assumption, and for the deduced maximum diameter of 10^{12} cm., it is found that the source in Cygnus emits radiation at a frequency of 80 Mc/s. as if it were a black-body at a temperature of 10^{14} deg. K.

§ 4. THE EVIDENCE FOR THE EXISTENCE OF HIGH-ENERGY PARTICLES IN THE SOURCE

The occurrence in the source of a true temperature of 10^{14} deg. K. would involve collisions by electrons having a mean energy of about 10^{10} electron volts. It is necessary to examine the possibility that special mechanisms may be responsible for the emission of intense radio waves by electrons of comparatively low energy.

The coherent oscillation of a large number of electrons might produce intense radio waves; both Martyn (1947) and Shklovsky (1947) have suggested that plasma oscillations in the solar corona could, in this way, account for the intense radiation associated with sunspots. This mechanism has been examined in some detail in a previous communication (Ryle 1949) which will be referred to as Paper 2. It was there concluded that the mechanism could not account for the emission of intense radio waves from the sun. It is now necessary to apply the arguments then used to determine whether plasma oscillations could account for the intense radiation from galactic sources.

When considering the possibility of plasma oscillations in the solar corona the range of electron density and magnetic field likely to occur in the region of the source were known from well-established observations. For the case of the galactic sources there is no information on the magnitude of the electron density or the magnetic field except that obtained from the radio-frequency observations.

It is first important to note that the radiation will only be able to reach the earth if the refractive index remains real at all points between the source and the earth. In the absence of a magnetic field the refractive index for a wave having a frequency of 80 Mc/s. vanishes at an electron density of about 8×10^7 cm^{-3} In the presence of an intense magnetic field the propagation of one circularly polarized component becomes possible for an appreciably greater electron density. If the source were situated in a region of great electron density and an intense magnetic field permitted the escape of one component, the radiation observed at the earth would, however, appear circularly polarized, a result which is not confirmed by experiment (Ryle and Smith 1948).

It is therefore concluded that the presence of an intense magnetic field is not essential to the emission of radio waves from the type of galactic source which has been described. (The variable component of the radiation from some galactic sources is suggestive of the radiation associated with sunspots, although preliminary experiments have suggested that this component is also randomly polarized.)

It might be suggested that circularly polarized radiation emitted from the two hemispheres of a star would be of opposite polarity and approximately equal intensity for an observer situated near the equatorial plane of the star. The observer would therefore deduce that the radiation was randomly polarized. Detailed observations of the polarization from both the intense galactic sources have, however, shown that the polarized component appears to be less than 5% of the total intensity; it is unlikely that both sources would be orientated with such precision as to give this kind of signal at the earth if the radiation were comprised of two circularly polarized waves of opposite polarity. It is therefore concluded that the observed random polarization is genuine; both components can therefore escape from the region of the source.

In the absence of an intense magnetic field in the region of generation, the radiation must be emitted at or above the level at which the electron density is about 8×10^7 cm^{-3}. Unless the temperature is less than about 10,000° K., ionization will be complete and the total particle density will not greatly exceed $1 \cdot 6 \times 10^8$ cm^{-3}.

It is thus apparent that the observed radiation can only be generated in regions where the electron density and magnetic field are similar to those in the solar corona. In Paper 2 reasons were given for supposing that plasma oscillations could not occur under the conditions of the solar corona, and it was suggested that the observed intensity of radio waves was due to the random motion of electrons having a mean energy corresponding to a temperature of 10^6 deg. K. in the normal corona and up to 10^{10} deg. K. near sunspots. The same arguments may now be applied to the radiation from galactic sources; in the absence of coherent electronic oscillations the observed intensity can only be accounted for by an electron temperature of at least 10^{14} deg. K.

The emission of radio waves from an electron gas and the maintenance of a large electron temperature in the solar corona by an electric field have been considered in an earlier paper (Ryle 1948); this paper will be referred to as Paper 1. It seems probable that similar mechanisms occur in the sources of intense galactic radiation; if an electric field is responsible for the maintenance of an electron temperature of 10^{14} deg. K., it is necessary that the energy gained by an electron between collisions shall be equal to the mean energy of thermal agitation of a gas at a temperature of 10^{14} deg. K. The electron density in the region of generation has been deduced from a determination of the refractive index at a frequency of

80 Mc/s.; the mean free path of an electron (for appreciable loss of energy) is then about 10^{10} cm. It may therefore be seen that an electric field of 1 volt cm^{-1} would be sufficient to maintain a mean electron energy of 10^{10} volts (corresponding to an electron temperature of 10^{14} deg. K.).

The existence of an electric field of this magnitude would be capable of accelerating charges of either sign to a very great energy and, as has been shown for the case of the solar corona (Paper 1), some of these particles would be able to escape from the stellar atmosphere with a large energy.

It is therefore suggested that the observed galactic sources are also capable of emitting particles having an energy corresponding to a potential of at least 10^{10} electron volts and that the total emission of particles from these bodies may be the origin of the primary cosmic rays.

The application of a theory of Alfvén (1937) to account for the maintenance of a large electron temperature in the solar corona has already been considered in Paper 1; the extension of these mechanisms to account for greater electron temperatures in the envelopes of other stellar bodies will now be considered.

§ 5. THE PRODUCTION OF HIGH-ENERGY PARTICLES IN A STELLAR ENVELOPE

Alfvén (1937) has shown that the potential developed between the poles and equator of a star of radius R_0 and polar field strength H_0 is of the order of $H_0R_0^2(\omega-\omega')$ where ω and ω' represent the angular velocities at the equator and the poles. For a star embedded in a region containing a large density of interstellar matter, the available potential may be significantly greater and is given by $H_0R_0^2\omega$. The recent observation (Babcock 1947 a) of stars having magnetic field strengths and rotational velocities very much greater than those of the sun therefore suggests that in certain stars a very large potential difference may be developed.

It was shown in Paper 1 that the intensity of radio waves emitted at any frequency by a stellar envelope depends on the mean electron energy in the region where the refractive index for that frequency vanishes. This energy depends on (*a*) the electric field and (*b*) the electron density, and the latter is determined by the frequency of the wave,* so that the intensity at a given frequency is determined solely by the magnitude of the electric field.

For a star in which the available potential is given by $H_0R_0^2\omega$, the electric field in the envelope will be of the order $H_0R_0\omega$. Observations by Babcock (1947 b) have shown that certain stars are likely to have, at the photosphere, a polar field strength of 5,000 gauss and a peripheral velocity ($R_0\omega$) of 2×10^7 cm. sec^{-1}. The product of these two quantities is about 10^4 times the corresponding figure for the sun; a star having these characteristics might therefore be expected to emit radio waves corresponding to a temperature about 10^4 times that observed for the sun. It is possible that stars having somewhat greater magnetic fields and rotational velocities may exist, and it is suggested that the mechanism described would be adequate to account for the maintenance of an electron temperature as great as 10^{14} deg. K. in the envelopes of certain stars.

§ 6. THE ANGULAR DISTRIBUTION OF THE SOURCES

Observations by Reber (1944) and Hey, Phillips and Parsons (1946) have shown that the intensity of radio waves is greatest from directions near the galactic equator, whereas the distribution of cosmic rays at the earth is highly isotropic. If the hypothesis put forward in this paper is correct, it is necessary to show how the observed isotropic distribution of cosmic rays could be produced during their passage from remote sources to the earth; it will be suggested that the important controlling factor is the magnetic fields of those stellar bodies which themselves produce the particles.

---

* This statement assumes that the emitted wave is unpolarized.

Babcock (1948) has shown that the magnetic field of certain stars could deflect a cosmic ray particle at a considerable distance; it is possible to calculate the "effective radius for large-angle scattering" for particles of different energies. If it is assumed that the stars responsible for scattering cosmic ray particles have, for particles of energy 10^{10} electron volts, an "effective scattering radius" R, and are distributed through the galaxy with density Δ, then if ρ is the effective radius of the galaxy, the probability P that a particle is deflected through a large angle as it traverses the galaxy is given by

$$P=\pi R^2\Delta\rho. \qquad \ldots\ldots(1)$$

If now it is assumed that certain stellar bodies are responsible for the production of the radio waves from the galaxy (and of the cosmic ray particles), and that these bodies also cause the deflection of the particles, it is possible to use the knowledge derived from the radio measurements to make an estimate of P.

The average radius of these stellar bodies for the emission of radio waves will be of the same order as the "photospheric" radius R_0; they will therefore subtend a total solid angle $\pi R_0^2\Delta\rho$ at the earth.

It has already been shown that the emission of radio waves at a frequency of 80 Mc/s. corresponds to an effective temperature of 10^{14} deg. K. at the source, so that if all the sources subtend a total solid angle $\pi R_0^2\Delta\rho$, the effective mean temperature measured at this frequency will be $\pi R_0^2\Delta\rho \times 10^{14}$ deg. K. Measurements at 80 Mc/s. have shown that the mean temperature of the galaxy is about 10^4 deg. K., and so it is deduced that $\pi R_0^2\Delta\rho \simeq 10^{-10}$.

Substitution from equation (1) now gives:

$$P=10^{-10}(R/R_0)^2. \qquad \ldots\ldots(2)$$

It would therefore be possible to deduce a value for P without a knowledge of Δ or ρ, providing $(R/R_0)^2$ were known.

From calculations based on Babcock's theory (1948), it can be shown that for a cosmic ray particle having an energy of 10^{10} electron volts, $(R/R_0)^2=3\times10^{-8}H_0R_0$ (where electromagnetic C.G.S. units are used). Reasons have already been given (§ 5) for supposing that galactic radio waves are produced in bodies for which $H_0R_0\omega$ is about 10^8 times the corresponding quantity for the sun. On the assumption that their angular velocity is the same as that of the sun, $H_0R_0\simeq 3\times10^{20}$ and $(R/R_0)^2\simeq 10^{13}$.

The probability that a cosmic ray particle of energy 10^{10} electron volts will be deflected through a large angle in passing through the galaxy would therefore be 1,000 (from equation (2)). Even if the angular velocities of the stellar bodies were 1,000 times as great as that of the sun, the probability of a violent deflection would be unity and the distribution of cosmic rays at the earth would still be isotropic.

It is therefore concluded that even in the absence of a general galactic magnetic field (Alfvén 1938, Babcock 1947 b) considerable scattering would be produced by the magnetic fields of the distributed sources themselves; the difference in angular distribution of cosmic rays and the radio waves from the galaxy is therefore not regarded as an argument against their generation in a common source.

§7. CONCLUSIONS

Observations of intense sources of radio-frequency radiation have suggested that certain stellar bodies are responsible for the emission, at metre wavelengths, of radiation whose intensity corresponds to the radiation from a black-body source at a temperature of 10^{14} deg. K. This estimate is based on deductions made from the intensity and its time fluctuations and on the assumption that the source is not nearer than the nearest known fixed star. So far it has not been possible to identify any of the sources with known stars, but it seems unlikely that a body having the characteristics deduced from the radio observations could pass unnoticed if it were at a comparatively small distance.

The possibility that intense radio waves could be emitted by the coherent motion of electrons of comparatively low energy has been examined, but the necessary conditions seem improbable in a stellar envelope. It is therefore concluded that the observations indicate the existence of regions in which the mean electron energy is of the order of 10^{10} electron volts. These regions would, under certain conditions, be capable of emitting particles having energies of this order of magnitude, and it is suggested that they are responsible for the production of primary cosmic rays.

The difficulty of identifying sources of such outstanding characteristics with those of visual observations may imply the existence of a type of star in which a high temperature corona is associated with a comparatively small visual temperature or opacity. This condition might arise in rapidly rotating stars having a large magnetic field.

The observations which have already been made on discrete sources of radio-frequency radiation suggest that it may be possible to explain the total radiation from the galaxy on the basis of a number of such sources. It therefore appears that a type of star which has so far not been classified in astronomy at visual wavelengths may be responsible for two major astrophysical phenomena—the origin of cosmic rays and of the intense radio waves from the galaxy.

The difference in the angular distribution of the radio waves and cosmic rays is accounted for by the deflecting effects of magnetic fields in the galaxy. In addition to previous theories of interstellar magnetic fields it is shown that the magnetic fields of the sources themselves are likely to cause an isotropic distribution of cosmic rays at the earth.

ACKNOWLEDGMENTS

I wish to acknowledge the great assistance provided by my colleagues at the Cavendish Laboratory, particularly by Mr. F. G. Smith, who has in addition carried out most of the analysis of our experimental records. I also owe much to stimulating discussions with Professor H. Alfvén.

The work was carried out as part of a programme supported at the Cavendish Laboratory by the Department of Scientific and Industrial Research.

REFERENCES

Alfvén, H., 1937, *Z. Phys.*, **105**, 633; 1938, *Phys. Rev.*, **54**, 97.

Babcock, H. W., 1947 a, *Astrophys. J.*, **105**, 105; 1947 b, *Phys. Rev.*, **72**, 83; 1948, *Ibid.*, **74**, 480.

BOLTON, J. G., 1948, *Nature, Lond.*, **162,** 141.
BOLTON, J. G., and STANLEY, G. J., 1948, *Nature, Lond.*, **161,** 312.
DAVIS, L., 1947, *Phys. Rev.*, **72,** 632.
HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W., 1946, *Nature, Lond.*, **158,** 234.
HEY, J. S., PHILLIPS, J. W., and PARSONS, S. J., 1946, *Nature, Lond.*, **157,** 296.
MARTYN, D. F., 1947, *Nature, Lond.*, **159,** 26.
REBER, G., 1944, *Astrophys. J.*, **100,** 279.
RYLE, M., 1948, *Proc. Roy. Soc.* A, **195,** 82 (referred to as Paper 1) ; 1949, *Proc. Phys. Soc.* A, **62,** (referred to as Paper 2).
RYLE, M., and SMITH, F. G., 1948, *Nature, Lond.*, **162,** 462.
SHKLOVSKY, J. S., 1947, *Nature, Lond.*, **159,** 752.
SWANN, W. F. G., 1933, *Phys. Rev.*, **43,** 217.

Pair-Production in Different Elements by γ-Rays

BY R. R. ROY
Wheatstone Physics Laboratory, King's College, University of London

Communicated by J. T. Randall ; MS. received 5*th October* 1948, *and in amended form* 6*th April* 1949

ABSTRACT. The production of electron pairs on bombardment of Pb, Pt, Sn, Ag and Cu by γ-rays from 200 mg. of Ra and its equilibrium products has been investigated. 1,200 photographs were taken for each element and the yield was 100 pairs in Pb, 85 in Pt, 72 in Sn, 63 in Ag and 50 in Cu. The distribution of energy of the positrons in Pb and Pt follows the shape of the theoretical curve of Bethe and Heitler, but the position of the maximum of the energy distribution curves in Sn, Ag and Cu shows a discrepancy with that to be expected theoretically. The theoretical difference between the average energies of the positrons and electrons in the elements studied, when plotted against Z, is consistently higher than the experimental, though both can be approximately represented by straight lines. The angular distribution of the positrons and electrons with respect to the incident γ-rays agrees with the theory.

§1. INTRODUCTION

THE creation of a pair of electrons of opposite charge in the Coulomb field of the nucleus by γ-rays of energy greater than $2m_0c^2$ is a well-known phenomenon, and the cross section of this process has been calculated by Bethe and Heitler (1934) and by Jaeger and Hulme (1936). The experimental cross section for pair-production in Pb has also been studied, in particular by Chadwick, Blackett and Occhialini (1934), and it was found to agree quite well with the theoretical cross section as given by Bethe and Heitler. The present experiments were undertaken to investigate (i) the energy spectra of the positrons of the pairs created in Pb, Pt, Sn, Ag and Cu, (ii) the difference between the average energy of the positrons and electrons as a function of the atomic number of the elements studied, (iii) the angle between the positron and electron and the angles of emission with respect to the direction of the incident γ-rays.

§2. EXPERIMENTAL ARRANGEMENTS

The experiments were carried out in a rubber-piston type cloud-chamber 16·5 cm. in diameter and 4 cm. in height; the standard source of illumination, viz. five 100-volt, 100-watt lamps, flashed across 200 volts, was used. A magnetic field of 400 gauss, supplied by a pair of Helmholtz coils, was used to determine the energy of the tracks. The measurements of the angles and curvatures of the tracks were taken after re-projecting the images from the two cameras. To reduce inaccuracy arising from multiple scattering, only those tracks whose radius of curvature lay between certain limits were taken into account. The source used was γ-rays from 200 mg. of Ra in equilibrium with its products, and, although the γ-rays from Ra have several lines above 1 MEV., in these experiments observations were restricted by energy consideration to those γ-rays the energy of which was 2·2 MEV. The foil of the element to be studied was placed in the cloud-chamber, against the glass ring; facing the side where the foil was situated was a lead block 45 cm. in length, 12 cm. in width, and 10 cm. in height, through which a hole 5 mm. in diameter had been drilled. This lead was necessary to prevent continuous ionization in the chamber due to γ-rays. At the other end of the block, away from the cloud-chamber, there was placed on one side of the hole a thin-walled iron container with the source inside it; on the other side, an electromagnet. When the expansion took place, the container was drawn towards the magnet; hence the source was in line with the hole and the γ-rays could pass through and bombard the element.

§3. RESULTS AND DISCUSSION

The results are given in Table 1.

Table 1

Element studied	Pb	Pt	Sn	Ag	Cu
Thickness of foil (μ)	15	15	20	20	25
No. of photographs taken	1200	1200	1200	1200	1200
No. of pairs found	100	85	72	63	50

Taking into account the density and thickness of the material, the yield of pairs bears out, approximately, the theoretical predictions.

(i) *The Energy Spectrum of the Positrons in Pb, Pt, Sn, Ag and Cu*

As mentioned above, the energy of the incident γ-rays was 2·2 MEV., and hence the energy available for the pair is 1·2 MEV. From the experimental results it was found that the energy spectra for Pb and Pt, and for Ag and Sn, were essentially the same; therefore, only the energy distribution curves for Pb and Sn (Figures 1 (*a*) and 1 (*b*)), and that for Cu (Figure 1 (*c*)), are given. They are plotted against the theoretical curve of Bethe and Heitler.

The abscissa in each figure represents the total energy of the positrons, that is, the sum of the kinetic and rest energies; the ordinate gives the ratio of the number of positrons at a particular energy to that of the total number of positrons observed from a given element. Examination of these curves shows that in the case of Pb and Pt the maximum lies at about 1·38 MEV., in the case of Sn and Ag at 1·25 MEV. and in the case of Cu at 1·15 MEV. It will be seen that for 2·2 MEV. incident γ-rays the energy distribution curve for Pb and Pt agrees in shape with the theoretical curve. However, whereas the energy at which the maximum of the theoretical curve occurs remains the same for all elements, the experimental results show a variation in the maximum from higher to lower values according as the element studied is heavier or lighter. In addition, it may be noted that the curve of Pb and Pt falls very rapidly above 1·4 MEV. and there are more particles at the higher than at the lower energies. The curve of Cu is almost symmetrical and the energy of the γ-quantum after the creation of a pair is more or less equally distributed between the positron and electron.

In order to calculate the energy of the positrons the curvatures of the tracks have to be measured. This measurement involves an inaccuracy at low energies, because of multiple scattering and the difficulty in ascertaining the small radii of curvature. At high energies inaccuracies are also liable to occur because of the large radius of curvature. The maximum error here was not more than 8% in the extreme cases. However, at those energies where the curves indicate the maximum yield, the measurement of the curvatures was accurate to within 4%. The discrepancy between experiment and theory noted in this region for lighter elements therefore seems real.

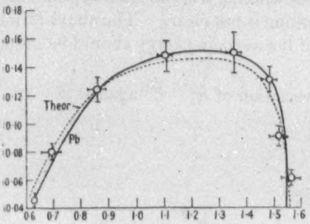


Figure 1 (*a*).

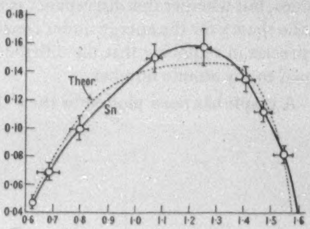


Figure 1 (*b*).

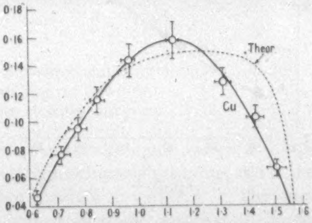


Figure 1 (*c*).

In one aspect the present experiment confirms that of Alichanow, Alichanian and Kosodaew (1936), who observed the pair-formation in Pb by 2·2 MeV. γ-rays from RaC and found that the maximum of the energy distribution curve of the positrons was at about 1·38 MeV.

(ii) *The Difference of Energy of Positron and Electron*

Approximate theoretical values for $\overline{E}_+ - \overline{E}_-$ can be calculated from the formula given by Bethe and Heitler:

$$\overline{E}_+ - \overline{E}_- = \frac{2m_0c^2Z}{137},$$

where Z is the atomic number.

Table 2 sets out the experimental results together with those given by the theory.

Table 2. $\overline{E}_+ - \overline{E}_-$

Element	Pb	Pt	Sn	Ag	Cu
Theory	0·6	0·57	0·36	0·34	0·21
Experiment	0·41 ± 0·06	0·38 ± 0·06	0·21 ± 0·05	0·19 ± 0·05	0·10 ± 0·04

The experimental results appear to be consistently lower than the theoretical values, but whether this discrepancy is real or whether it arises from approximation in the theory for the energy under consideration is not clear. The above formula is precise in requiring that the difference of the average energy should be proportional to the atomic number.

A graph has been plotted to show the variation of $E_+ - \overline{E}_-$ against Z.

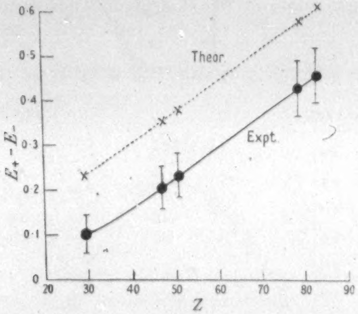


Figure 2.

Figure 2 shows the graphs of both theoretical and experimental results. The theoretical results show a straight line with a slope $h = 0{\cdot}0073$. The experimental curve is almost a straight line having a slope $h = 0{\cdot}0063 \pm 0{\cdot}001$.

(iii) *The Angular Distribution*

According to Bethe and Heitler, the fraction of the particles projected between the angles θ and $\theta + d\theta$ to the incident photon is given by

$$\frac{\theta\, d\theta}{(\Theta^2 + \theta^2)^2}$$

where $\Theta = \mathbf{m_0 c^2}/E$, where E is the total energy of either of the projected particles, i.e. the sum of the kinetic and rest energies. The maximum of the theoretical distribution of either positrons or electrons for γ-rays of 2·2 MEV. when plotted against the direction of the incident photon lies at about 15°. From the experimental data it is found that the angular distribution of both positrons and electrons is the same in each element, and that the maximum of the curves of Pb, Pt, Sn, Ag and Cu lies at about 18°. Therefore only one curve has been used to illustrate the results.

Figure 3 shows the angular distribution for the positrons in Pb together with theoretical curve, given by the dotted line. The experimental curve follows the general shape of the theoretical distribution, and to this extent indicates fair agreement between the theoretical and experimental results.

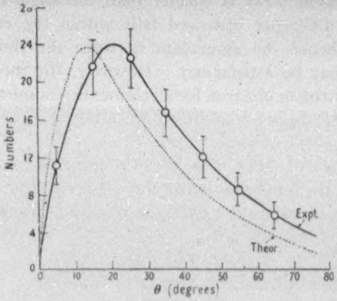


Figure 3.

The measurements of the angles between the directions of emission of the positrons and electrons are given in Table 3.

Table 3

Element	Pb	Pt	Sn	Ag	Cu
Average angle (°) between positron and electron	31	30	30	28	28
Average angle (°) between positron and γ-ray	13	13	14	13	13
Average angle (°) between electron and γ-ray	18	17	16	15	15

The Table shows that the average angle between the positron and electron in the different elements is approximately the same, and is about 30°. The measurements of the angles are accurate to within 3°. The mean of the average angles between the electron and γ-ray is 16°, and that of the positron and γ-ray is 13°.

Klarmann and Bothe (1936) investigated the problem of pair-production in krypton and xenon by ThC″ γ-rays of energy 2·6 Mev. and found that the average angle between the positron and electron increased with the atomic number. Thus for krypton it was 19° and for xenon 42°, but the number of pairs studied was only 13. The present observation does not confirm this, but agrees with Simóns and Zuber (1937), who found that the average angle in argon was 30° and in iodine 32°. The photograph (see Plate) shows the formation of a pair in copper by a γ-ray of energy 2·2 Mev.

§ 4. CONCLUSION

Lawson (1949) has discussed the limitations of the theory of Bethe and Heitler, and it appears that this theory is not suitable for precise comparison with the present experimental results. For γ-rays of energy less than 3 Mev. Lawson states that the calculation of Jaeger and Hulme gives a correction of opposite sign to that required when comparing the experimental results with the theory of Bethe and Heitler. Since with heavy elements, however, the lack of agreement between theory and experiment possibly arises from the inapplicability of the Born approximation, while for light elements it may arise from the use of the

Fermi–Thomas atom, satisfactory comparison between experiment and any present theory is hardly possible. Bethe and Heitler's theory predicts that the angular distribution should be the same for both electrons and positrons. The experimental results show that the angle between the direction of emission of positron and incident γ-ray is smaller than the angle between the electron and γ-ray, but the difference observed falls within the error involved in the measurements and hence the agreement between the theory and experiment in this connection may be satisfactory. However, the most probable angle of ejection of either positron or electron for all elements is somewhat higher, as shown in Figure 3, than the theory indicates. In addition, the difference of average energy between the positron and electron is consistently lower than that to be expected from the theory. The change in the position of the maximum of the energy spectrum of the positrons with the element studied is quite distinct, and cannot be explained within the limits of probable error.

ACKNOWLEDGMENT

The author wishes to express his thanks to Dr. F. C. Champion for his continued interest in the progress of the experiment and for his helpful discussions.

REFERENCES

ALICHANOW, A. I., ALICHANIAN, A. E., and KOSODAEW, M. S., 1936, *J. Phys. Radium*, **7**, 163.
BETHE, H., and HEITLER, W., 1934, *Proc. Roy. Soc.* A, **146**, 83.
CHADWICK, J., BLACKETT, P. M. S., and OCCHIALINI, G. P. S., 1934, *Proc. Roy. Soc.* A, **144**, 235.
JAEGER, J. C., and HULME, H. R., 1936, *Proc. Roy. Soc.* A, **153**, 443.
KLARMANN, H., and BOTHE, W., 1936, *Z. Phys.*, **101**, 489.
LAWSON, J. L., 1949, *Phys. Rev.*, **75**, 433.
SIMONS, L., and ZUBER, K., 1937, *Proc. Roy. Soc.* A, **159**, 383.

Structure and Diamagnetic Anisotropy of P-Benzoquinodimethane in Connection with those of P-Benzoquinone

By M. G. EVANS*, J. DE HEER† AND J. GERGELY‡,

Communicated by C. A. Coulson ; MS. received 1st *February* 1949

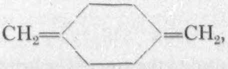
ABSTRACT. The interest in the diamagnetic properties of quinodimethane lies in the fact that this molecule can be considered as a prototype for quinone. The diamagnetic anisotropy, $\Delta\chi$, of p-benzoquinodimethane has been calculated by the molecular orbital method following the treatment given by F. London in 1937. We find :

$$\Delta\chi_{\text{p-benzoquinodimethane}} \sim 0{\cdot}3\Delta\chi_{\text{benzene}}$$

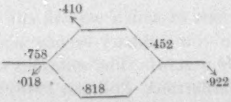
Calculations of the diamagnetic anisotropy of quinone itself fail because the perturbation used here breaks down in this case, in which all the centres in the molecule are not identical.

The result obtained for the quinodimethane can be discussed in connection with the observations by K. Lonsdale and K. S. Krishnan on quinone. We conclude that although $\Delta\chi$ of quinone is much reduced compared with that for benzene, we need not postulate, to account for this observation, a complete or nearly complete localization of the π-electrons in this structure.

THOUGH Szwarc (1947) has only recently shown the presence of p-quinodimethane,



in the products of the pyrolysis of p-xylene, this compound has been, for several reasons, studied theoretically since 1945 (Diatkina, Namiot, Syrkin 1945, Coulson, Craig, Maccoll, Pullman 1947, Pullman, Berthier, Pullman 1948). In the diagram below we show the free valencies and bond-orders§ as calculated by Pullman *et al.* (1948) by means of the molecular orbital method :



For reasons which will be discussed below, we were interested in the diamagnetic anisotropy, $\Delta\chi$, of this compound. A theoretical approach is the only possible one at present. Unfortunately so far no well-established general theory of diamagnetic anisotropy has been put forward. F. London, in 1937, developed a method for calculating $\Delta\chi$ in certain aromatic systems with reference to $\Delta\chi$ for benzene as standard. This method is essentially equivalent to a molecular orbital treatment of the mobile π-electrons in these systems with the usual L.C.A.O. approximation.† By choosing variation functions of a suitable type, the energies ϵ^1 of the molecular orbitals, in a magnetic field of strength H in the direction perpendicular to the plane of the molecule, are obtained as the solutions of a secular equation of the general form

---

* Department of Chemistry, The University, Manchester.

† Department of Chemistry, The University, Manchester; Ramsay Memorial Fellow, the Netherlands.

‡ Laboratory of Physical Biology, National Institute of Health, Bethesda, U.S.A.

§ For the definition of these quantities see e.g. Coulson (1947).

$$|\mathscr{H}_{kl}-\epsilon^1\delta_{kl}|=0. \qquad \ldots\ldots(1)$$

$\mathscr{H}_{kl}$ reduces to the ordinary Coulomb integral, α, for $k=l$. $\mathscr{H}_{kl}=\beta\exp(2\pi if_{kl})=\beta\eta_{kl}$ if k and l are adjacent atoms (β is the ordinary resonance integral) and $\mathscr{H}_{kl}$ is zero otherwise. δ_{kl} is the Kronecker symbol and the quantity f_{kl} above expresses in suitable units the magnetic flux through the triangle formed by the centre of the molecule and the bond between the two atoms k and l. In the case of the p-quinodimethane molecule an eighth-order equation is obtained, which can be reduced by means of group-theory (the only symmetry operation allowed in the presence of the magnetic field being an inversion with respect to the centre of the molecule) to the following two:

$$\begin{vmatrix} x & \beta & 0 & 0 \\ \beta & x & \eta_s & \eta_s^* \\ 0 & \eta_s^* & x & \eta_d \\ 0 & \eta_s & \eta_d^* & x \end{vmatrix}=0 \quad \ldots\ldots(2a) \quad \text{and} \quad \begin{vmatrix} x & \beta & 0 & 0 \\ \beta & x & \eta_s & -\eta_s^* \\ 0 & \eta_s^* & x & \eta_d \\ 0 & -\eta_s & \eta_d^* & x \end{vmatrix}=0, \quad \ldots\ldots(2b)$$

where x is written for $(\alpha-\epsilon^1)/\beta$, the η^*s are the complex conjugates of the ηs, and the suffixes s and d indicate a "single" and a "double" bond in the ring of the molecule. Write

$$\eta_s^2\eta_d=\exp(\pi if), \qquad \ldots\ldots(3)$$

where $f[(\boldsymbol{e}/\boldsymbol{hc})H]^{-1}$ is put equal to the area of the benzene ring. In this approximate treatment we can neglect the small difference between the areas of the benzene and the "quinoid" ring in quinodimethane. As f is a small quantity, we obtain from (2*a*), (2*b*):

$$x^4-4x^2+2x+1=(\pi f)^2x, \qquad \ldots\ldots(4a)$$

$$x^4-4x^2-2x+1=-(\pi f)^2x. \qquad \ldots\ldots(4b)$$

These equations, the solutions of which we will call x^1, reduce to the ones we would have obtained from the ordinary secular equations for p-benzoquinodimethane if f were zero (H is zero). The solutions of the latter, which we will call x_0, are known (Pullman, Berthier, Pullman 1948) and as f is a small quantity we obtain from (4*a*), (4*b*)

$$(x^1-x_0)/(\pi f)^2=\frac{x_0}{4x_0^3-8x_0+2}, \qquad \ldots\ldots(5a)$$

$$(x^1-x_0)/(\pi f)^2=\frac{-x_0}{4x_0^3-8x_0-2}. \qquad \ldots\ldots(5b)$$

From (5*a*), (5*b*) the x^1s and so the ϵ^1s are readily calculated, and by summing these over all the occupied orbitals we obtain the total mobile energy E^1 in the presence of the magnetic field in terms of $(\pi f)^2$. Finally, following London (1937), the diamagnetic anisotropy, $\Delta\chi$, is assumed to be equal to the diamagnetic susceptibility due to the "motion" of the mobile electrons in the plane of the molecule, and thus is given by

---

† A summary of the approximations involved in this method is given by Coulson and Dewar (1947).

$$\Delta\chi = -\frac{1}{H}\frac{\partial E}{\partial H}. \qquad \ldots\ldots(6)$$

The result of our calculation (which is expressed most simply in terms of $\Delta\chi_{\text{benzene}}$), is $\Delta\chi_{\text{p-benzoquinodimethane}} = 0{\cdot}27\Delta\chi_{\text{benzene}}$.

This value is surprisingly low if we compare it with the experimental data for p-benzoquinone (such data lacking for the quinodimethane itself). Lonsdale (1939, see also Lonsdale and Krishnan 1936) reported a value for $\Delta\chi_{\text{p-benzoquinone}}$ of 40·6 compared with 61·9 for benzene. This relatively low diamagnetic anisotropy has been put forward by one of us (Evans 1946) as an argument to approximate a quinone as a molecule in which the π-electrons are completely localized in the "double" bonds. On the other hand McMurry's investigations of the spectra of C=O compounds (1941) indicate that in a quinone π-electrons must occupy molecular orbitals extending all through the molecule just as in the quinodimethanes. Recently the authors have confirmed the latter point of view (Evans, Gergely, de Heer 1949). We have shown that there still is an appreciable delocalization of π-electrons in a quinone structure even when we introduce in the secular equations appropriate parameters to take into account the extra binding energy of the C=O bond as compared with the C=C bond and the extra electronegativity of the O-atom as compared with that of the C-atom. But now the problem remained to account theoretically for the low diamagnetic anisotropy of these structures. We tried to calculate the $\Delta\chi$ for p-benzoquinone by introducing similar parameters to those mentioned above in the secular equations (2). In this way, however, we found a "negative diamagnetic anisotropy" (or a "paramagnetic contribution" of the π-electrons), which of course has no meaning at all and at present only seems to show the limitations of the London method.

The fact that this method breaks down for quinones is not surprising, as here a second-order effect is calculated by means of a first-order perturbation calculus. The fairly good results nevertheless obtained by London might be correlated with the fact that he only considered molecules which were described as "crossable hydrocarbons" by Coulson and Rushbrooke (1940) and as "alternant hydrocarbons" by Coulson and Longuet-Higgins (1947). These authors showed that only for these hydrocarbons is the molecular orbital method in its simplest form self-consistent. So, while this method is still useful to calculate quantities such as mobile energy, bond-orders and free valencies in both non-crossable hydrocarbons and hetero molecules (in the absence of an external field), it might easily give completely wrong results if used to calculate a second-order effect such as the diamagnetic anisotropy.

As a result of this difficulty we had to confine ourselves to the calculation of $\Delta\chi$ for p-benzoquinodimethane and to consider this molecule as the prototype of a quinone (as has been done before (Pullman, Berthier and Pullman 1948, Diatkina and Syrkin 1946)). As we saw, this computation gave us a very low $\Delta\chi$, though the relatively high resonance energy of $1{\cdot}924\,\beta$ indicates an appreciable delocalization, and the bond-orders shown in the diagram above indicate that the ring in this molecule approaches more closely to a benzene ring than to a completely localized quinonoid structure.

Although, remembering the approximate character of the London method, we do not attach any undue importance to the numerical value we have found, we tentatively suggest that a relatively small perturbation of the "pure benzene structure" in an aromatic compound can cause a rapid decrease in the diamagnetic anisotropy. While a satisfactory explanation of this fact cannot be given at present, we may conclude conversely that the low diamagnetic anisotropy found experimentally in p-benzoquinone need not be correlated at all with a complete, or nearly complete, localization of the double bonds in this molecule.

Note added in proof. It is easy to modify our calculations so as to include the effect of an overlap integral S between the atomic wave functions on adjacent atoms. As is well known, the inclusion of this integral modifies the solutions, x_0, of the unperturbed secular equation by a factor $1/(1-Sx_0)$. Brooks (1940) has shown that in a magnetic field the solutions (x^1-x_0) of equations (5 a, b) above have to be "normalized" by $1/(1-Sx_0)^2$, the diamagnetic properties being far more sensitive to this non-orthogonality correction than the energy. Although with $S=0{\cdot}25$ the contributions of the individual electrons, ϵ^1, change appreciably, the total mobile energy E^1 remains practically unchanged.

Our final result now becomes

$$\Delta\chi_{\text{p-benzoquinodimethane}} = 0{\cdot}25\Delta\chi_{\text{benzene}}$$

ACKNOWLEDGMENT

We wish to thank Professor C. A. Coulson for many helpful discussions.

REFERENCES

Brooks, Harvey., 1941, *J. Chem. Phys.*, **9,** 463.

Coulson, C. A., 1947, *The Labile Molecule*: Discussion of the Faraday Soc., p. 9.

Coulson, C. A., Craig, D. P., Maccoll, A., and Pullman, A., 1947, *The Labile Molecule*: Discussion of the Faraday Soc., p. 36.

Coulson, C. A., and Dewar, M. J. S., 1947, *The Labile Molecule*: Discussion of the Faraday Soc., p. 54.

Coulson, C. A., and Longuet-Higgins, H. C., 1947, *Proc. Roy. Soc.* A, **192,** 16.

Coulson, C. A., and Rushbrooke, G. S., 1940, *Proc. Camb. Phil. Soc.*, **36,** 193.

Diatkina, M., Namiot, A. J., and Syrkin, J., 1945, *C.R. Acad. Sci. U.R.S.S.*, **48,** 233.

Diatkina, M., and Syrkin, J., 1946, *Acta Phys. Chem. U.R.S.S.*, XXI, **5,** 921.

Evans, M. G., 1946, *Trans. Faraday Soc.*, **42,** 113.

Evans, M. G., Gergely, J., and de Heer, J., 1949, *Trans. Faraday Soc.*, **45,** 312.

London, F., 1937, *J. Phys. Radium*, **8,** 397.

Lonsdale, K., 1939, *Proc. Roy. Soc.* A, **171,** 559.

Lonsdale, K., and Krishnan, K. S., 1936, *Proc. Roy. Soc.* A, **156,** 597.

McMurry, H. L., 1941, *J. Chem. Phys.*, **9,** 241.

Pullman, A., Berthier, G., and Pullman, B., 1948, *Bull. Soc. Chim. France*, 5^e ser., **15,** 450.

Szwarc, M., 1947, *The Labile Molecule*: Discussion of the Faraday Soc., p. 46.

The Computation of Wave Functions in Momentum Space—I: The Helium Atom

By R. McWEENY* and C. A. COULSON†

MS. received 19*th April* 1949

ABSTRACT. The possibilities of obtaining precise momentum wave functions for atoms by direct solution of the wave equation in momentum space are examined in some detail. An iterative method of approximating to the wave function is employed in computing the momentum distribution function for the helium atom. Although in this instance considerable accuracy is achieved, formidable difficulties arise in extending the calculations to more complicated atoms.

§1. INTRODUCTION

FOR some time the importance of momentum wave functions has been appreciated: in particular, such functions are fundamental to any discussion of the Compton scattering of x rays (DuMond 1933) and the inelastic scattering of fast electrons (Hughes and Mann 1938).

Hitherto the possibility of computing momentum wave functions directly, by solving the eigenvalue equation in the momentum representation, has usually been dismissed as impracticable and the required functions have been obtained by an application of transformation theory to the more familiar coordinate space wave functions (Duncanson and Coulson 1945). Although this procedure is satisfactory when the approximate solution is built up solely from hydrogen-like orbitals with appropriate screening constants, it proves impossible to transform to momentum space the most accurate wave functions in which electron correlation is adequately represented, the integrals involved being quite insuperable. This difficulty has prompted a further examination of the possibilities of a direct solution in momentum space. The methods of calculation are essentially those employed by Svartholm (1945) in the momentum space treatment of nuclear systems: a distinction arises, however, in that emphasis will be placed upon the computation of the wave functions themselves rather than the associated energies; for this purpose the iterative method is particularly suitable. In the present paper, the n-electron problem is formulated in momentum space and in the case of the helium atom the iterative method of solution is applied with considerable success. In more general problems the "screening constant" wave functions of Slater and others may be improved to some extent by iteration but there appears to be little prospect, at present, of obtaining from such improved functions the necessary momentum distribution function.

§2. FORMULATION OF THE PROBLEM

For generality we consider a system of N particles, the first N–1 being electrons of mass m and charge $-e$, while the Nth is a nucleus of mass M and charge $+Ze$. Since it is impossible to avoid a certain amount of transformation theory a notation for matrix elements and transformation functions must be introduced: the Dirac notation seems most convenient for the formal derivation of the eigenvalue

---

* University College, Oxford, now at King's College, University of Durham.
† Wheatstone Physics Laboratory, King's College, London.

equation. For practical purposes however, a familiarity with representation theory is not essential since we may begin our calculations starting with equation (8). In what follows we shall use either of the following equivalent notations for the momentum wave function and its complex conjugate:

$$\langle p_{1x} p_{1y} \ldots p_{Nz} | \rangle \equiv \Phi(p_{1x} \ldots p_{Nz}),$$
$$\langle | p_{1x} p_{1y} \ldots p_{Nz} \rangle \equiv \Phi^{*}(p_{1x} \ldots p_{Nz}).$$

The set of commuting observables $p_{1x}, \ldots p_{Nz}$ may suggest that Φ is essentially a function of the Cartesian components of momentum : the choice of coordinates in which the wave function is expressed is, of course, arbitrary and the dependence of Φ on the momenta may be abbreviated to $\Phi = \Phi(\mathbf{p}_1 \ldots \mathbf{p}_N)$ i.e. a function of N momenta, each momentum being dependent upon three variables. In many cases the spherical polar coordinates (p, θ, ϕ) of each momentum prove most convenient. Often the whole set of momenta will be abbreviated to $\mathbf{p}$. The formalism should be clear from the following equivalences:

$$\langle \mathbf{p} | \rangle \equiv \Phi(\mathbf{p}) \equiv \Phi(\mathbf{p}_1 \mathbf{p}_2 \ldots \mathbf{p}_N) \equiv \Phi(p_1, \theta_1, \phi_1, p_2, \theta_2, \phi_2, \ldots \phi_N),$$
$$d\mathbf{p} \equiv d\mathbf{p}_1 \, d\mathbf{p}_2 \ldots d\mathbf{p}_N \equiv p_1^2 \sin\theta_1 \, dp_1 \, d\theta_1 \, d\phi_1 \, p_2^2 \ldots d\phi_N.$$

In momentum space the eigenvalue equation for the energy becomes

$$\left[\frac{1}{2M} \mathbf{p}_N^2 + \frac{1}{2m} \sum_{j=1}^{N-1} \mathbf{p}_j^2\right] \langle \mathbf{p} | \rangle + \int \langle \mathbf{p} | V | \mathbf{p}' \rangle \, d\mathbf{p}' \langle \mathbf{p}' | \rangle = E \langle \mathbf{p} | \rangle. \qquad \ldots\ldots(1)$$

The representative $\langle \mathbf{p} | V | \mathbf{p}' \rangle$ of the potential energy operator is most expediently found by transformation theory. In coordinate space

$$V = \sum_{j=1}^{N-1} V_{jN}(|\mathbf{r}_{jN}|) + \sum_{j<k}^{N-1} V_{jk}(|\mathbf{r}_{jk}|), \qquad \ldots\ldots(2)$$

$\mathbf{r}_{jk}$, $\mathbf{r}_{jN}$ being the vector separations of particles (j, k), (j, N). A single term may be transformed to momentum space according to

$$\langle \mathbf{p}_1 \ldots \mathbf{p}_N | V_{jk} | \mathbf{p}_1' \ldots \mathbf{p}_N' \rangle = \int \ldots \int \langle \mathbf{p}_1 | \mathbf{r}_1 \rangle \ldots \langle \mathbf{p}_N | \mathbf{r}_N \rangle V(|\overline{\mathbf{r}_j - \mathbf{r}_k}|)$$
$$\times \, d\mathbf{r}_1 \ldots d\mathbf{r}_N \langle \mathbf{r}_1 | \mathbf{p}_1' \rangle \ldots \langle \mathbf{r}_N | \mathbf{p}_N' \rangle, \qquad \ldots\ldots(3)$$

where $\mathbf{r}_1 \ldots \mathbf{r}_N$ are the position vectors of the N particles, and $d\mathbf{r}_1 \ldots d\mathbf{r}_N$ are volume elements in three dimensions. The transformation functions are

$$\langle \mathbf{p} | \mathbf{r} \rangle = h^{-3/2} \exp\{-i \mathbf{p} . \mathbf{r} / \hbar\}, \quad \langle \mathbf{r} | \mathbf{p} \rangle = h^{-3/2} \exp\{i \mathbf{p} . \mathbf{r} / \hbar\}.$$

The multiple integral (3) may be separated into a product of integrals over the 3-dimensional subspaces and finally an introduction of relative coordinates gives (cf. Svartholm 1945)

$$\langle \mathbf{p}_1 \ldots \mathbf{p}_N | V_{jk} | \mathbf{p}_1' \ldots \mathbf{p}_N' \rangle$$
$$= \delta(\overline{\mathbf{p}_j - \mathbf{p}_j'} + \overline{\mathbf{p}_k - \mathbf{p}_k'}) \prod_{r \neq j \neq k} \delta(\overline{\mathbf{p}_r - \mathbf{p}_r'}) . \langle \mathbf{p}_j | V_{jk} | \mathbf{p}_j' \rangle, \qquad \ldots\ldots(4)$$

where $$\langle \mathbf{p}_j | V_{jk} | \mathbf{p}_j' \rangle = h^{-3} \int \exp\{-i \overline{\mathbf{p}_j - \mathbf{p}_j'} . \mathbf{r} / \hbar\} V_{jk}(|\mathbf{r}|) \, d\mathbf{r}. \qquad \ldots\ldots(5)$$

With attractions of Coulombic form the appropriate integral is easily evaluated *, giving

$$\langle \mathbf{p}_j | \frac{1}{|\mathbf{r}|} | \mathbf{p}_j' \rangle = \frac{1}{2\pi^2 \hbar} \cdot \frac{1}{|\mathbf{p}_j - \mathbf{p}_j'|^2}. \qquad \ldots\ldots(6)$$

---

* A convergence factor must be employed.

Substitution of this result in (1) gives an integral equation for the eigenvalues and eigenfunctions. We confine ourselves to solutions in which there is no resultant translatory motion of the whole system by setting

$$\langle \mathbf{p}_1, \mathbf{p}_2, \ldots\ldots \mathbf{p}_N|\rangle = \delta(\mathbf{p}_1+\mathbf{p}_2+\ldots\ldots+\mathbf{p}_N)\langle \mathbf{p}_1\,\mathbf{p}_2\ldots\ldots \mathbf{p}_{N-1}|\rangle.$$

Putting $p_0^2=-2mE$, $R=m/M$, $\lambda=m/\pi^2\hbar$ we find finally

$$\left[p_0^2+2R\sum_{j<k}^{N-1}\mathbf{p}_j\cdot\mathbf{p}_k+(1+R)\sum_{j=1}^{N-1}p_j^2\right]\langle \mathbf{p}_1\ldots\ldots \mathbf{p}_{N-1}|\rangle$$
$$=-\lambda\left[\sum_{j=1}^{N-1}\int\langle \mathbf{p}_j|V_{jN}|\mathbf{p}_j'\rangle d\mathbf{p}_j'\langle \mathbf{p}_1\ldots\ldots \mathbf{p}_j'\ldots\ldots \mathbf{p}_{N-1}|\rangle\right.$$
$$\left.+\sum_{j<k}^{N-1}\int\langle \mathbf{p}_j|V_{jk}|\mathbf{p}_j'\rangle d\mathbf{p}_j'\langle \mathbf{p}_1\ldots\ldots \mathbf{p}_j'\ldots\ldots \overline{\mathbf{p}_j+\mathbf{p}_k-\mathbf{p}_j'}\ldots\ldots \mathbf{p}_{N-1}|\rangle\right]. \quad \ldots\ldots(7)$$

The accuracy of solution does not, in most cases, warrant the inclusion of R: iteration with a finite value of R would introduce a correction for the motion of the nucleus. We shall regard R as vanishingly small, and, with a slight change of variables, write the eigenvalue equation for an atomic system, for which the forces are Coulombic, in the following form, where $n=N-1$ is the number of electrons in the atom:

$$\sum_0^n p_j^2\langle \mathbf{p}_1\ldots\ldots \mathbf{p}_n|\rangle=\lambda\left[Z\sum_{j=1}^n\int\frac{d\mathbf{p}}{|\mathbf{p}|^2}\langle \mathbf{p}_1\ldots\ldots \overline{\mathbf{p}_j-\mathbf{p}}\ldots\ldots \mathbf{p}_n|\rangle\right.$$
$$\left.-\sum_{j<k}^n\int\frac{d\mathbf{p}}{|\mathbf{p}|^2}\langle \mathbf{p}_1\ldots\ldots \overline{\mathbf{p}_j-\mathbf{p}}\ldots\ldots \overline{\mathbf{p}_k+\mathbf{p}}\ldots\ldots \mathbf{p}_n|\rangle\right. \quad \ldots\ldots(8)$$

§3. METHOD OF SOLUTION

The integral equation (8) may be written quite generally as

$$T(\mathbf{p})\langle \mathbf{p}|\rangle=\lambda\int\langle \mathbf{p}|V|\mathbf{p}'\rangle d\mathbf{p}'\langle \mathbf{p}'|\rangle. \quad \ldots\ldots(9)$$

In momentum space $T(\mathbf{p})$ is always algebraic, being in fact a homogeneous quadratic form in the momenta,

$$T(\mathbf{p}_1\mathbf{p}_2\ldots\ldots \mathbf{p}_n)=(p_0^2+p_1^2+p_2^2+\ldots\ldots p_n^2).$$

Solutions of (9) usually exist for a discrete set of negative values of E (positive values of p_0^2): we are primarily interested in the (algebraically) least of these. For the purposes of solution it is convenient to regard λ as the eigenvalue parameter: ultimately we determine an expression for an upper bound to the least eigenvalue λ^0 in terms of the constants of the problem and the unknown p_0^2: on giving λ its true value $1/\pi^2$ a corresponding value of p_0^2 may be calculated. It is easily seen that the value of E obtained in this way is an upper bound to the least energy eigenvalue.

The most powerful method of obtaining approximate solutions of an equation of this kind appears to be that contained in a theorem originally due to Kellogg (1922). A simple proof is given by Svartholm (1945), who shows that the solution of (9) may be approached by iteration in the following way.

Choosing an arbitrary initial function $\langle \mathbf{p}|\rangle^0$ form the iterated functions $\langle \mathbf{p}|\rangle^1$, $\langle \mathbf{p}|\rangle^2, \ldots\ldots$ according to the rule $\langle \mathbf{p}|\rangle^{j+1}=\dfrac{1}{T(\mathbf{p})}\displaystyle\int\langle \mathbf{p}|V|\mathbf{p}'\rangle d\mathbf{p}'\langle \mathbf{p}'|\rangle^j$, and the integrals

$$T_j=\int{}^j\langle|\mathbf{p}\rangle T(\mathbf{p})d\mathbf{p}\langle \mathbf{p}|\rangle^j,$$
$$W_j=\iint{}^j\langle|\mathbf{p}\rangle d\mathbf{p}\langle \mathbf{p}|V|\mathbf{p}'\rangle d\mathbf{p}'\langle \mathbf{p}'|\rangle^j.$$

Next form the ratios $\lambda_j = T_j/W_j$, $\lambda_{j+\frac{1}{2}} = W_j/T_{j+1} (j=0, 1, \ldots.)$. Then the sequence $\lambda_0 > \lambda_{\frac{1}{2}} > \lambda_1 > \ldots.$ converges monotonically to the numerically least eigenvalue (λ^0) of (9). This result is dependent on the positive definiteness of the operators T and V.

For our purposes the trend of the iterated functions themselves is more important. It is not difficult to show that

$$\langle \mathbf{p}|\rangle^j \rightarrow \frac{\text{const.}}{(\lambda^0)^j} \langle \mathbf{p}|0\rangle, \quad j \rightarrow \infty,$$

where $\langle \mathbf{p}|0\rangle$ is the eigenfunction corresponding to the eigenvalue λ^0.

For, writing $\langle \mathbf{p}|\rangle^0 = \Sigma_s \langle \mathbf{p}|s\rangle\langle s|\rangle^0$, where $|s\rangle$ is the sth eigenstate of (9), and making use of the definition $T(\mathbf{p})\langle \mathbf{p}|s\rangle = \lambda^s \int \langle \mathbf{p}|V|\mathbf{p}'\rangle d\mathbf{p}' \langle \mathbf{p}'|s\rangle$, we have

$$\langle \mathbf{p}|\rangle^1 = \sum_s \frac{1}{T(\mathbf{p})} \int \langle \mathbf{p}|V|\mathbf{p}'\rangle d\mathbf{p}' \langle \mathbf{p}'|s\rangle\langle s|\rangle^0 = \sum_s \frac{1}{\lambda^s} \langle \mathbf{p}|s\rangle\langle s|\rangle^0.$$

Repetition then yields $\langle \mathbf{p}|\rangle^j = \sum\limits_s \left(\frac{1}{\lambda^s}\right)^j \langle \mathbf{p}|s\rangle\langle s|\rangle^0$, and since λ^0 is by hypothesis the numerically least eigenvalue, the desired result follows immediately.

The rapidity of convergence of the process thus depends essentially on the separation of the eigenvalues $\lambda^0 \lambda^1 \ldots. \lambda^s \ldots.$ (higher eigenvalues of λ corresponding to excited states of a hypothetical system rather than of the actual one) and is difficult to assess precisely. It appears, however, that with a carefully chosen initial function one iteration is often sufficient to produce a rather accurate approximation to the eigenfunction.

Even at this stage an advantage peculiar to the iterative method is evident. Although in many cases neither T_1 nor W_1 is calculable, the formal expression for the eigenfunction with which they are associated, i.e. the eigenfunction corresponding to the third approximation, λ_1, to the eigenvalue, has already occurred in the calculation of W_0. The superiority of this approach over a variational method, in which the approximate eigenvalue must first be evaluated and secondly minimized in order to assess the best values of the parameters, is apparent.

§4. APPLICATION TO THE HELIUM ATOM

(i) *Choice of Initial Function.*

With $n=2$ the eigenvalue equation becomes, using atomic units,

$$(p_0^2 + p_1^2 + p_2^2)\,\Phi(\mathbf{p}_1\mathbf{p}_2) = \lambda[Z(I_1(\mathbf{p}_1\mathbf{p}_2) + I_2(\mathbf{p}_1\mathbf{p}_2)) - I_{12}(\mathbf{p}_1\mathbf{p}_2)], \quad \ldots\ldots(10)$$

where

$$\left.\begin{aligned} I_1 &= \int \frac{d\mathbf{p}}{|\mathbf{p}|^2} \Phi(\overline{\mathbf{p}_1 - \mathbf{p}}, \mathbf{p}_2), \quad I_2 = \int \frac{d\mathbf{p}}{|\mathbf{p}|^2} \Phi(\mathbf{p}_1, \overline{\mathbf{p}_2 - \mathbf{p}}) \\ I_{12} &= \int \frac{d\mathbf{p}}{|\mathbf{p}|^2} \Phi(\overline{\mathbf{p}_1 - \mathbf{p}}, \overline{\mathbf{p}_2 + \mathbf{p}}), \quad \lambda = 1/\pi^2. \end{aligned}\right\} \quad \ldots\ldots(11)$$

In constructing a first approximation to the wave function for an atom it is customary to regard the mutual interaction of the electrons as a negligible perturbation. Neglect of all integrals such as I_{12} then permits a separation of (10), or more generally (8), and indicates an approximate solution of the product form analogous to the Hartree type solution in coordinate space. When spin functions are included and the required antisymmetry is introduced Φ is most generally a linear superposition of determinantal functions such as

$$\Phi_1 = \begin{vmatrix} \phi_1(\mathbf{p}_1)\alpha(1), & \phi_1(\mathbf{p}_1)\beta(1), & \phi_2(\mathbf{p}_1)\alpha(1)\ldots \\ \phi_1(\mathbf{p}_2)\alpha(2), & \phi_1(\mathbf{p}_2)\beta(2), & \phi_2(\mathbf{p}_2)\alpha(2)\ldots \\ \ldots & \ldots & \ldots \end{vmatrix} \qquad \ldots\ldots(12)$$

where $\phi_1(\mathbf{p})$, $\Phi_2(\mathbf{p})$, ... are the various possible wave functions for one electron in the field of the nucleus.

In the ground state problem at present under consideration

$$\phi_1(\mathbf{p}) = \phi_2(\mathbf{p}) = \phi_{1s}(\mathbf{p})$$

and for a Hamiltonian not including spin variables, the antisymmetric spin factor in the wave function may, of course, be omitted. We may then write

$$\Phi^0(\mathbf{p}_1\mathbf{p}_2) = \phi_{1s}(\mathbf{p}_1).\ \phi_{1s}(\mathbf{p}_2).$$

As in coordinate space the single particle functions are of simple analytical form: in order to introduce some flexibility into the functions we admit a variable parameter and write

$$\Phi^0(\mathbf{p}_1\mathbf{p}_2) = \frac{1}{(p_1^2+a^2)^2(p_2^2+a^2)^2}, \qquad \ldots\ldots(13)$$

where, using atomic units, a is numerically equal to an "effective" nuclear charge. There is no need to normalize the function at this stage.

Since this function is simply the Fourier transform of the screening constant wave function in coordinate space, we expect, and in fact obtain, an identical variational value of the energy from a calculation of λ_0. The evaluation of the integrals T_0, W_0 (or their equivalents). has been carried through independently by Copson (1943): it is worth noting that the calculation is not appreciably more difficult than the corresponding calculation in coordinate space.

In view of our previous observations as to the relative merits of variation and iteration, (13) appears to be the most suitable initial function, further refinement being most conveniently introduced by iteration.

(ii) *The Iterated Function and Variation Treatment.*

Using the initial function (13) the first iterated function is given by

$$(p_0^2+p_1^2+p_2^2)\Phi^1(\mathbf{p}_1\mathbf{p}_2) = [Z(I_1(\mathbf{p}_1\mathbf{p}_2)+I_2(\mathbf{p}_1\mathbf{p}_2)-I_{12}(\mathbf{p}_1\mathbf{p}_2)], \qquad \ldots\ldots(14)$$

where, from (11), we find easily

$$I_1(\mathbf{p}_1\mathbf{p}_2) = \frac{\pi^2}{a}\frac{1}{(p_1^2+a^2)(p_2^2+a^2)^2}, \quad I_2(\mathbf{p}_1\mathbf{p}_2) = \frac{\pi^2}{a}\frac{1}{(p_1^2+a^2)^2(p_2^2+a^2)},$$

$$I_{12}(\mathbf{p}_1\mathbf{p}_2) = \int \frac{d\mathbf{p}}{|\mathbf{p}|^2}\frac{1}{\{(\mathbf{p}_1-\mathbf{p})^2+a^2\}^2\{(\mathbf{p}_2+\mathbf{p})^2+a^2\}^2}. \qquad \ldots\ldots(15)$$

W_0 may now be calculated:

$$W_0 = \int\int \Phi^{0*}(p_0^2+p_1^2+p_2^2)\Phi^1 d\mathbf{p}_1\, d\mathbf{p}_2.$$

By performing the integrations over $\mathbf{p}_1$ and $\mathbf{p}_2$ first the direct evaluation of I_{12} is circumvented and we obtain

$$W_0 = \frac{\pi^6}{16a^9}(Z - 5/16). \qquad \ldots\ldots(16)$$

The calculation of T_0 is straightforward

$$T_0 = \frac{\pi^4}{64a^8}\left(2 + \frac{p_0^2}{a^2}\right). \qquad \ldots\ldots(17)$$

So $\lambda_0 = \dfrac{1}{4a\pi^2} \cdot \dfrac{2a^2 + p_0^2}{Z - 5/16}$ possesses a least value $\dfrac{p_0}{\pi^2\sqrt{2}}\,\dfrac{1}{(Z-5/16)}$ when $a^2 = p_0^2/2$.

Now, according to (11) we must put this minimum value equal to $1/\pi^2$. This gives

$$p_0^2 = 2(Z - 5/16)^2 \text{A.U.}, \; E = 2(Z - 5/16)^2 W_{\mathrm{H}}. \qquad \ldots\ldots(18)$$

In this way the most suitable value of a is obtained. In the actual computation of the iterated function we shall therefore take $a = 2 - 5/16 = 1{\cdot}6875$ A.U.

Before proceeding with the computation a further remark upon the choice of numerical parameters is necessary, for (16) involves not only a but p_0^2.

Now in practice we cannot make an infinite number of iterations and so are never able to compute exactly the appropriate λ in the relation $\lambda = \lambda(p_0^2)$. This implies that we can never exactly infer the inverse relation $p_0^2 = p_0^2(\lambda)$ and consequently can never compute the true value of p_0^2 in the case of interest where $\lambda = 1/\pi^2$. Fortunately, however, the iterated wave function shows only a weak dependence upon p_0^2 and the use of a nearly correct p_0^2 instead of the completely correct one has no significant effect on the resultant wave function. In the present case we shall use the accurately known value $p_0^2 = 5{\cdot}8074$ A.U.

The significance of the various members of the iterated function becomes apparent on giving p_0^2 its variational value, $a^2/2$. With this somewhat inferior choice we find

$$\frac{I_1(\mathbf{p}_1\mathbf{p}_2) + I_2(\mathbf{p}_1\mathbf{p}_2)}{(p^0 + p_1^2 + p_2^2)} \quad \rightarrow \quad \frac{\pi^2}{a}\,\frac{1}{(p_1^2 + a^2)^2(p_2^2 + a^2)^2},$$

which is simply a multiple of the initial function. Accordingly we rewrite (14) in the form

$$\Phi^1(\mathbf{p}_1\mathbf{p}_2) = I'(\mathbf{p}_1\mathbf{p}_2) - I'_{12}(\mathbf{p}_1\mathbf{p}_2), \qquad \ldots\ldots(19)$$

where

$$I'(\mathbf{p}_1\mathbf{p}_2) = \frac{Z(I_1(\mathbf{p}_1\mathbf{p}_2) + I_2(\mathbf{p}_1\mathbf{p}_2))}{(p_0^2 + p_1^2 + p_2^2)},$$

$$I'_{12}(\mathbf{p}_1\mathbf{p}_2) = \frac{I_{12}(\mathbf{p}_1\mathbf{p}_2)}{(p_0^2 + p_1^2 + p_2^2)}.$$

Here $I'(\mathbf{p}_1\mathbf{p}_2)$ in itself represents a slightly improved wave function when p_0^2 is given its true value: in this case

$$I'(\mathbf{p}_1\mathbf{p}_2) = \frac{Z\pi^2}{a}\,\frac{p_1^2 + p_2^2 + 2a^2}{(p_0^2 + p_1^2 + p_2^2)(p_1^2 + a^2)^2(p_2^2 + a^2)^2}. \qquad \ldots\ldots(20)$$

which is no longer a simple product, $f(\mathbf{p}_1)\cdot f(\mathbf{p}_2)$. In this way the probability distribution of $\mathbf{p}_1$ is made to depend on the instantaneous value of $|\mathbf{p}_2|$: a similar small correction arises from I'_{12}, as may be verified by expanding the denominator of the integrand in terms of $\cos(\mathbf{p}_1, \mathbf{p})$, $\cos(\mathbf{p}_2, \mathbf{p})$: all terms beyond the first are dependent on the mutual inclination of the electron momenta as well as on their magnitudes.

(iii) *The Momentum Distribution Function.*

Although the function $\Phi^1(\mathbf{p}_1\mathbf{p}_2)$ is of some intrinsic interest it is of little practical value: since the electrons are indistinguishable the only observable quantity is the probability of any electron having a given momentum $\mathbf{p}$. This latter function is obtained by integrating the probability density over the momentum coordinates of all electrons but one. In the present case it is accordingly

$$P^1(\mathbf{p}_1)=\int\Phi^{1*}(\mathbf{p}_1\mathbf{p}_2)\,\Phi^1(\mathbf{p}_1\mathbf{p}_2)d\mathbf{p}_2. \qquad \ldots\ldots(21)$$

Moreover, since we invariably conduct our experiments on an array of randomly oriented atoms or molecules rather than on a single system, a further integration over the solid angle ω_1 about $\mathbf{p}_1$ is necessary. The quantity of greatest importance in the present case is thus

$$P(p_1)=\int\!\!\int\Phi'^*(\mathbf{p}_1\mathbf{p}_2)\Phi'(\mathbf{p}_1\mathbf{p}_2)\,d\mathbf{p}_2p_1^2\,d\omega_1. \qquad \ldots\ldots(22)$$

$P(p_1)\,dp_1$ then measures the probability of observing a momentum with magnitude between p_1 and p_1+dp_1.

When simple product wave functions are employed the integrations are easily performed since the integrals separate into a product of "one-electron" integrals (cf. Duncanson and Coulson 1945). In particular, we record the normalized probability distribution corresponding to the initial variation function:

$$P^0(p)=\frac{32a^5p^2}{\pi(p^2+a^2)^4}.$$

With a better wave function, however, such as $\Phi^1(\mathbf{p}_1\mathbf{p}_2)$, a separation of this kind is not possible and the integration over all $\mathbf{p}_2$-space may constitute a major difficulty.

Neglecting terms which are of the second order of small quantities

$$P'(\mathbf{p}_1)=\int(I'^2-2I''\,I'_{12})\,d\mathbf{p}_2, \qquad \ldots\ldots(23)$$

where I'' denotes the approximation to I' which results on taking $\mathbf{p}_0^2=a^2/2$; this choice of p_0^2 in I' makes a quite negligible change in the small correction term $2I'I'_{12}$ and at the same time greatly simplifies the computation.

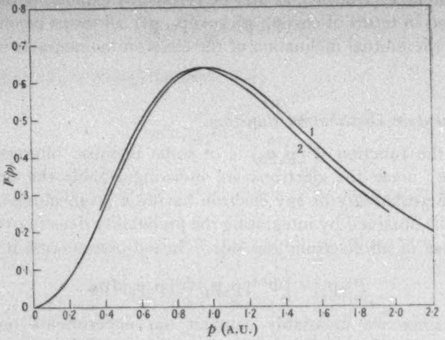
It is of interest to plot the corrected distribution retaining only the major term in the integral. We find easily

$$\int I'^2d\mathbf{p}_2=\frac{2\pi^5Z^2}{a^2}\left[\frac{I(2/22,p_1)}{(p_1^2+a^2)^4}+\frac{2I(2/32,p_1)}{(p_1^2+a^2)^3}+\frac{I(2/42,p_1)}{(p_1^2+a^2)^2}\right], \qquad \ldots\ldots(24)$$

where

$$I(2/r2,p_1)=\int_{-\infty}^{+\infty}\frac{z^2\,dz}{(z^2+a^2)^r(z^2+p_0^2+p_1^2)^2}.$$

The various integrals have been tabulated as functions of p_1 and the modified distribution curve plotted: to the scale adopted in the Figure the curve is, however, almost indistinguishable from the first approximation.



1. Initial approximation $P^0(p)$. 2. Corrected function $P^1(p)$.

The contribution of the second term in the integrand of (23) is

$$2\int I''I'_{12}d\mathbf{p}_2$$

$$=\frac{2\pi^2 Z}{a}\iint\frac{1}{(p_1^2+a^2)^2(p_2^2+a^2)^2}\cdot\frac{1}{(p_0^2+p_1^2+p_2^2)}\cdot\frac{d\mathbf{p}\,d\mathbf{p}_2}{|\mathbf{p}|^2\{(\mathbf{p}_1+\mathbf{p})^2+a^2\}^2\{(\mathbf{p}_2-\mathbf{p})^2+a^2\}^2}.$$

We may write

$$\int I''I'_{12}d\mathbf{p}_2=\frac{\pi^2 Z}{a}\frac{1}{(p_1^2+a^2)^2}\int\frac{K'(\mathbf{p})\,d\mathbf{p}}{|\mathbf{p}|^2\{(\mathbf{p}_1+\mathbf{p})^2+a^2\}^2},\qquad\ldots\ldots(25)$$

where

$$K'(\mathbf{p})=\int\frac{1}{(p_0^2+p_1^2+p_2^2)}\,\frac{1}{(p_2^2+a^2)}\,\frac{d\mathbf{p}_2}{\{(\mathbf{p}_2-\mathbf{p})^2+a^2\}^2}.$$

It is convenient to evaluate first

$$K(\mathbf{p})=\int\frac{1}{(p_2^2+c^2)(p_2^2+b^2)}\,\frac{d\mathbf{p}_2}{\{(\mathbf{p}_2-\mathbf{p})^2+a^2\}^2},\ [b^2=p_0^2+p_1^2],\qquad\ldots\ldots(26)$$

from which $K'(\mathbf{p})$ may be obtained by differentiating with respect to c and subsequently letting $c\to a$.

Taking polar coordinates about $\mathbf{p}$

$$K(\mathbf{p})=\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty}\frac{p_2^2dp_2\,\sin\theta\,d\theta\,d\phi}{(p_2^2+c^2)(p_2^2+b^2)(p_2^2+p^2+a^2-2pp_2\cos\theta)^2}.$$

The angle integrations are easily accomplished and the denominator of the resultant integral may be factorized to give

$$K(\mathbf{p})$$

$$=4\pi\int_0^{\infty}\frac{z^2\,dz}{(z+ic)(z-ic)(z+ib)(z-ib)(\overline{z-p}+ia)(\overline{z-p}-ia)(\overline{z+p}+ia)(\overline{z+p}-ia)}.$$

$$\ldots\ldots(27)$$

By symmetry it is sufficient to evaluate the integral between the limits ±∞, this being accomplished by calculus of residues on completing the contour by an infinite semicircle in the upper half-plane.

We proceed directly to the calculation of $K'(\mathbf{p})$, writing down the residues, differentiating with respect to c, and letting $c \to a$. After some reduction we obtain

$$K'(\mathbf{p}) = \frac{\pi^2}{a(a^2-b^2)^2}\left[\frac{Ap^4+Bp^2+C}{(p^2+4a^2)^2(p^4+2Pp^2+Q^2)}\right], \qquad \ldots\ldots(28)$$

where

$$A=(a-b)^2,\ B=(11a^4-16a^3b+2a^2b^2+3b^4),$$
$$C=(10a^6-32a^5b+30a^4b^2-10a^2b^4+2b^6),$$
$$P=(a^2+b^2),\ Q=(a^2-b^2).$$

The second integration, that indicated in (25), is rather more involved: it may be similarly reduced to a contour integral resembling (27) but with additional poles. The final result is

$$\int I'' I'_{12}\, d\mathbf{p}_2 = \frac{\pi^4}{2a^2(p_1^2+a^2)^2}[R_1+R_2+R_3], \qquad \ldots\ldots(29)$$

where

$$R_1 = \frac{2(B-Aa^2)}{LM} + \frac{4(16Aa^4-4Ba^2+C)(a^2-p_1^2)}{LM^2} + \frac{(16Aa^4-4Ba^2+C)(p_1^4-26a^2p_1^2+57a^4)}{4a^2L^2M},$$

$$R_2 = \frac{AX+BY+CZ}{2L^2(p_1^2+a^2)^5},$$

$$R_3 = \frac{a}{(a+b)[2ab(p_1^2+a^2)-(a^4-b^4)]},$$

and

$$L = p_1^4+10a^2p_1^2+9a^4,$$
$$M = p_1^4-6a^2p_1^2-7a^4,$$
$$X = (p_1^2+a^2)^3(p_1^6+7a^2p_1^4+27a^4p_1^2+21a^6),$$
$$Y = (p_1^2+a^2)^2(p_1^6-a^2p_1^4-5a^4p_1^2-3a^6),$$
$$Z = (p_1^8-12a^2p_1^6-42a^4p_1^4-44a^6p_1^2-15a^8).$$

Remembering $b=\sqrt{(p_0^2+p_1^2)} \simeq \sqrt{(2a^2+p_1^2)}$ to the present approximation, we may tabulate the above expressions as functions of p_1. Since the result is angle independent the integration over the solid angle ω_1 yields only a factor 4π. The final momentum distribution including this further correction is illustrated in the Figure. The curve indicates a considerable decrease in the mean momentum, as would be expected from the virial theorem. For in momentum space

$$T = \frac{1}{2m}\int_0^\infty P(p)\, p^2\, dp$$

and a shift of the peak of $P(p)$ towards regions of lower momentum implies a shift of weight from large to small values of p, with a consequent reduction in $\overline{T}$. Normalization is not easily effected analytically and for purposes of comparison the curve is plotted to the same peak height as with the first approximation.

(iv) *Possibilities of Further Applications.*

We conclude with an enquiry into the possibilities of extending the analysis just developed for helium to other atomic systems. Since the eigenvalue equation whose solution we are seeking refers essentially to a monatomic system, and association of any kind is known to have a profound effect on the momentum distribution, the next simplest atom to be dealt with should, logically, be neon, intervening atoms in the periodic table being incapable of sustained independent existence.

It is essential that the initial function should possess the completely antisymmetric form demanded by the Pauli Principle, for otherwise iteration would lead to a "collapse" of the initial functions into the lowest available orbitals. This naturally introduces enormous complications but there is still no formal difficulty in writing down an iterated function. Thus we find

$$\Phi^1(\mathbf{p}_1 \ldots . \mathbf{p}_n) = \frac{S}{(p_0^2+p_1^2+ \ldots . +p_n^2)} - \frac{S_{rs}}{(p_0^2+p_1^2+ \ldots . +p_n^2)},$$

where S denotes the appropriate antisymmetric sum of a number of terms containing only one-electron integrals (every one of which is integrable with a screening constant initial function) and S_{rs} denotes a similar sum of "correlation integrals" (such as I_{12} of the last section).

As already noted, the leading term in the expansion of each correlation integral could be evaluated and with the inclusion of such terms a rather accurate wave function might be obtained.

In this way correlation effects would be represented to a considerable extent, and the corrected wave function would no longer be a simple product of one-electron functions. Unfortunately, however, a serious difficulty is simultaneously introduced. No function is of very great practical utility unless it permits the computation of $P(p)$: the suggested function after iteration does not lend itself to such a computation, the task of integrating over all the momenta but one proving a formidable obstacle to further progress. For that reason we have not proceeded with the calculation of the iterated function.

REFERENCES

Copson, E. T., 1943, *Proc. Roy. Soc. Edinb.*, **61**, 260.

DuMond, J. W. M., 1933, *Rev. Mod. Phys.*, **5**, 1.

Duncanson, W. E., and Coulson, C. A., 1945, *Proc. Phys. Soc.*, **57**, 190.

Hughes, A. L., and Mann, M. M., 1938, *Phys. Rev.*, **53**, 50.

Kellogg, O. D., 1922, *Math. Ann.*, **86**, 14.

Svartholm, N. V., 1945, *The Binding Energies of the Lightest Atomic Nuclei* (Stockholm: Lund).

The Computation of Wave Functions in Momentum Space— II: The Hydrogen Molecule Ion

By R. McWEENY *
University College, Oxford

Communicated by C. A. Coulson ; MS. received 19th April 1949

ABSTRACT. A new approach is made to the problem of computing accurate momentum distribution functions for electrons in molecules. The usual molecular orbital wave function (linear combination of atomic orbitals) has already been employed for this purpose but has proved inadequate in attempts to explain the width of the Compton profile. Using the method of a previous paper, the l.c.a.o. type of molecular orbital is formally corrected by iteration. In most cases it seems likely that the corrected orbital will be associated with a greater spread of the momentum distribution and will thus lead to a broader Compton profile : the broadening should become more marked with increasing delocalization of the electron and may apparently be very considerable. A complete calculation has confirmed these conclusions in the case of the hydrogen molecule ion : a preliminary treatment of the neutral molecule is given.

§1. INTRODUCTION

In Part I† of this work the distribution of momentum in a simple monatomic system was computed directly by approximate solution of the eigenvalue integral equation in momentum space ; certain features of the iterative method employed were found particularly suited to this purpose.

The object of the present investigation is to determine whether similar methods may be employed in computing the momentum distribution of electrons in molecules. A complete calculation is at present confined to the simplest of all diatomic molecules, the hydrogen molecule ion, but certain general effects of aggregation upon individual atomic orbitals are predicted. Results suggest that the use of more accurate wave functions might give very much better agreement than hitherto between the computed and observed profiles for Compton scattering from molecules and solid scatterers.

§2. THE MOTION OF AN ELECTRON ABOUT *n* CENTRES OF FORCE

(i) *Formulation of the Problem*

For generality let us consider the motion of an electron about *n* centres of force located at points with position vectors $\mathbf{R}_1, \mathbf{R}_2, \ldots \mathbf{R}_n$. Let the position vector of the electron be $\mathbf{r}$. We shall require the momentum space representative of the potential energy operator: this may be obtained using the methods and notation of I. If V_s be the potential function for the *s*th centre

$$V=\Sigma_s V_s(|\mathbf{r}-\mathbf{R}_s|).$$

A single term is transformed to momentum space according to

$$\langle \mathbf{p}|\,V_s(|\mathbf{r}-\mathbf{R}_s|)\,|\mathbf{p}'\rangle = h^{-3}\int \exp\{-i\,\overline{\mathbf{p}-\mathbf{p}'}\,.\,\mathbf{r}/\hbar\}\,.\,V_s(|\mathbf{r}-\mathbf{R}_s|)\,d\mathbf{r}.$$

---

* Now at King's College, University of Durham.

† Page 509 of this Journal : hereafter referred to as I.

With a change of variable we easily find, using atomic units as in I,

$$\langle \mathbf{p}|\, V_s(|\mathbf{r}-\mathbf{R}_s|)\,|\,\mathbf{p}'\rangle = \exp\{-i\,\overline{\mathbf{p}-\mathbf{p}}'\,.\,\mathbf{R}_s\}\,.\,\langle \mathbf{p}|\, V_s(|\mathbf{r}|)\,|\,\mathbf{p}'\rangle, \qquad \ldots\ldots(1)$$

where $\langle \mathbf{p}|\, V_s(|\mathbf{r}|)\,|\,\mathbf{p}'\rangle$ is simply the momentum space transform of V_s with the sth centre regarded as origin.

The eigenvalue equation may thus be written

$$(p_0^2+p^2)\phi(\mathbf{p}) = -2m \sum_{s=1}^{n} \int \exp\{-i\,\overline{\mathbf{p}-\mathbf{p}}'\,.\,\mathbf{R}_s\}.\langle \mathbf{p}|\, V_s\,|\,\mathbf{p}'\rangle\, d\mathbf{p}'\phi(\mathbf{p}'). \qquad \ldots\ldots(2)$$

If we suppose the electron to be in the vicinity of one particular nucleus and assume that all other terms in the potential function are therefore negligible compared with, say, V_s, it is easily verified that

$$\phi(\mathbf{p}) = \exp(-i\mathbf{p}\,.\,\mathbf{R}_s)\,.\,\phi_s(\mathbf{p}) \qquad \ldots\ldots(3)$$

is a solution of (2) provided $\phi_s(\mathbf{p})$ satisfies

$$(p_0^2+p^2)\phi_s(\mathbf{p}) = -2m\int \langle \mathbf{p}|\, V_s\,|\,\mathbf{p}'\rangle\, d\mathbf{p}'\phi_s(\mathbf{p}').$$

The function $\phi_s(\mathbf{p})$ is thus simply an eigenfunction for the motion of the electron about the sth nucleus alone, situated at the origin. We may therefore call $\exp(-i\mathbf{p}\,.\,\mathbf{R}_s)$ a "displacement operator". Putting

$$D_s(\mathbf{p}) = \exp(-i\mathbf{p}\,.\,\mathbf{R}_s), \qquad D_s^{-1}(\mathbf{p}) = \exp(+i\mathbf{p}\,.\,\mathbf{R}_s),$$

the displacement of a physical system from **0** to **R** induces the following transformation of associated operators and wave function

$$\langle \mathbf{p}|\, V\,|\,\mathbf{p}'\rangle \to D_s(\mathbf{p})\langle \mathbf{p}|\, V\,|\,\mathbf{p}'\rangle D_s^{-1}(\mathbf{p}'),$$
$$\phi(\mathbf{p}) \to D_s(\mathbf{p})\,.\,\phi(\mathbf{p}).$$

The eigenvalue equation (2) may thus be written

$$(p_0^2+p^2)\phi(\mathbf{p}) = -2m \sum_{s=1}^{n} \int D_s(\mathbf{p})\langle \mathbf{p}|\, V_s\,|\,\mathbf{p}'\rangle D_s^{-1}(\mathbf{p}')\, d\mathbf{p}'\phi(\mathbf{p}'). \qquad \ldots\ldots(2')$$

A superposition wave function precisely equivalent to the molecular orbital in coordinate space may be formed from terms such as (3). We shall call such a function $\phi^0(\mathbf{p})$ and take it as our initial function, namely

$$\phi^0(\mathbf{p}) = \sigma_1 D_1(\mathbf{p})\,.\,\phi_1(\mathbf{p}) + \sigma_2 D_2(\mathbf{p})\,.\,\phi_2(\mathbf{p}) + \ldots \sigma_n D_n(\mathbf{p})\,.\,\phi_n(\mathbf{p}). \qquad \ldots\ldots(4)$$

$\phi_1(\mathbf{p}) \ldots \phi_n(\mathbf{p})$ are the momentum space transforms of atomic orbitals about centres 1, ... n, and will in general be approximated by hydrogen-like wave functions. $\sigma_1 \ldots \sigma_n$ are best assessed by the usual molecular orbital treatment.

(ii) *The Iterative Correction*

Taking (4) as the starting point of our treatment the first iterated function $\phi^1(\mathbf{p})$ may immediately be written down. From (2′)

$$\begin{aligned}\phi^1(\mathbf{p}) &= \frac{-2m}{(p_0^2+p^2)} \sum_s \sum_t \sigma_t \int D_s(\mathbf{p})\langle \mathbf{p}|\, V_s\,|\,\mathbf{p}'\rangle D_s^{-1}(\mathbf{p}')\, d\mathbf{p}' D_t(\mathbf{p}')\phi_t(\mathbf{p}') \\ &= \frac{-2m}{(p_0^2+p^2)} \sum_s \sum_t \sigma_t I^{st} D_s(\mathbf{p}),\end{aligned} \qquad \ldots\ldots(5)$$

where

$$I^{st} = \int \langle \mathbf{p}|\, V_s\,|\,\mathbf{p}'\rangle \exp\{-i\,\mathbf{p}'\,.\,\overline{\mathbf{R}_t-\mathbf{R}_s}\}\, d\mathbf{p}'\phi_t(\mathbf{p}'). \qquad \ldots\ldots(6)$$

Even at this stage inferences as to the nature of the corrections may be made. A typical term, $I^{st}D_s(\mathbf{p})$, of the summand in (5) may be said to represent the perturbing influence of the *s*th nucleus on the momentum distribution of an electron moving in the field of the *t*th nucleus alone. Even with neglect of all terms except those for which $s=t$ a correction results, for, to this approximation

$$\phi^1(\mathbf{p}) = \frac{-2m}{(p_0^2+p^2)} \sum_{t=1}^{n} \sigma_t D_t(\mathbf{p}) \int \langle \mathbf{p} | V_t | \mathbf{p}' \rangle d\mathbf{p}' \phi_t(\mathbf{p}'). \qquad \ldots\ldots(7)$$

But, by definition of the atomic orbital $\phi_t(\mathbf{p})$

$$[(p_0^2)_t + p^2]\phi_t(\mathbf{p}) = -2m \int \langle \mathbf{p} | V_t | \mathbf{p}' \rangle d\mathbf{p}' \phi_t(\mathbf{p}'), \qquad \ldots\ldots(8)$$

so that

$$\phi^1(\mathbf{p}) = \sum_{=1}^{n} \sigma_t D_t(\mathbf{p}) \frac{(p_0^2)_t + p^2}{p_0^2 + p^2} \phi_t(\mathbf{p}). \qquad \ldots\ldots(9)$$

Now (9) is simply a linear superposition molecular orbital in which the individual atomic orbitals have been modified by a factor which is an angle-independent function of $\mathbf{p}$. Moreover, we can at once predict the effect of this correction on the momentum distribution function, for $(p_0^2)_t$ and p_0^2 may be regarded as, at least, approximately known quantities. Quite generally we may suppose $p_0^2 > (p_0^2)_t$ since we are concerned solely with the ground state in which the electron occupies a bonding molecular orbital. Each atomic orbital in (9) is thus multiplied by a monotonically increasing function of p: consequently the resultant momentum distribution is a less rapidly decreasing function of p than the initial approximation would suggest. To a first approximation the half-width of the momentum distribution function is in this way always increased by the perturbing effect of neighbouring nuclei. Since, generally speaking, the difference between $(p_0^2)_t$ and p_0^2 becomes more marked as the binding of the electron to the *t*th centre becomes less, we may say that the increase in half-width of the distribution function over that predicted on a simple l.c.a.o. treatment (Coulson 1941) will become greater as localization of the electron at a particular centre decreases. This result is in agreement with experimental results on Compton scattering. Although the width of the Compton profile is not simply proportional to the width of the momentum distribution there is always a marked similarity between the two curves and it may at least be shown that an increase in width of one implies an increase in width of the other. The computed profile agrees very well with experimental curves for monatomic scatterers but for molecular and solid scatterers the computed width, using the l.c.a.o. approximation to the wave function, falls far short of the observed width (Duncanson and Coulson 1941, p. 417). Assuming the approximate additivity of one-electron distributions (Coulson 1941), these considerations suggest that refinement of the wave functions may lead to much closer agreement. This result is substantiated in a subsequent section.

The modification of the initial wave function so far discussed may be described as a spherically symmetrical distortion of each atomic orbital in the initial linear combination. The modifications imposed by those terms so far neglected are generally smaller and introduce an angle-dependent distortion, or polarization, of each atomic orbital, together with a "mixing" of the orbitals. Reference to (5) shows that we may still regard the molecular orbital as a linear combination of atomic orbitals provided $\Sigma_t \sigma_t I^{st}$ be interpreted as a distorted atomic orbital about the *s*th centre. In addition to the dominant correction already discussed there will then be a contribution from every other centre, *t*, dependent on three factors: (i) the form of the orbital about the *t*th centre, (ii) the relative importance of this orbital, measured by σ_t, (iii) the magnitude and direction of the position vector of *t* relative to *s*.

In this way each atomic orbital takes up, on correction, the symmetry of the whole molecule. It is difficult to assess generally the effect of this correction on the momentum distribution function; calculation shows that in the case of the hydrogen molecule ion it is to oppose the correction first discussed. The polarization effect will, however, become less marked with increasing nuclear separation and with increasing symmetry of the configuration, features characteristic of the solid state.

§ 3. THE HYDROGEN MOLECULE ION

For a homonuclear system such as H_2^+ with nuclei 1 and 2 at $-\mathbf{R}$ and $+\mathbf{R}$, (2) and (4) assume particularly simple forms. With a nuclear charge Z and attractions of Coulombic form we find, using atomic units

$$(p_0^2+p^2)\phi(\mathbf{p}) = \frac{2Z}{\pi^2}\int \frac{\cos(\overline{\mathbf{p}-\mathbf{p}'}\,.\,\mathbf{R})}{|\mathbf{p}-\mathbf{p}'|^2}\,d\mathbf{p}'\phi(\mathbf{p}'), \qquad \ldots\ldots(10)$$

$$\phi^0(\mathbf{p}) = 2\cos(\mathbf{p}\,.\,\mathbf{R})\,.\,\phi_1(\mathbf{p}), \qquad \ldots\ldots(11)$$

where $\phi_1(\mathbf{p})$ is an atomic orbital for motion about either centre and symmetry has imposed the condition $\sigma_1=\sigma_2$.

The momentum distribution corresponding to a wave function (11) has been obtained elsewhere (Coulson 1941). For purposes of comparison the normalized curve is plotted in Figure 1 as $P^0(p)$.

As we have already noted (I) a variational method of approximating to the wave function, using a linear function system such as

$$\phi_v(\mathbf{p}) = \sum_{m,n}\frac{P_m(\cos\theta)}{(p^2+a^2)^n},$$

is often precluded by the fact that it necessitates a computation of the energy eigenvalue and this may be vastly more difficult in momentum space than in coordinate space. For this reason we adopt the iterative approach.

The first iterated function, corresponding to (5), is most conveniently written

$$\phi^1(\mathbf{p}) = \frac{1}{p_0^2+p^2}\,[I_s(\mathbf{p})+I_p(\mathbf{p})], \qquad \ldots\ldots(12)$$

where $I_s(\mathbf{p}) = 2\cos(\mathbf{p}\,.\,\mathbf{R})\int\frac{d\mathbf{p}}{|\mathbf{p}-\mathbf{p}'|^2}\,\phi_1(\mathbf{p}'), \qquad \ldots\ldots(13)$

and $I_p(\mathbf{p}) = 2\,\mathrm{Re}\,[\exp(i\mathbf{p}\,.\,\mathbf{R})]\int\frac{d\mathbf{p}'}{|\mathbf{p}-\mathbf{p}'|^2}\exp(-2i\mathbf{p}'\,.\,\mathbf{R})\,.\,\phi_1(\mathbf{p}'). \qquad \ldots\ldots(14)$

$I_s(\mathbf{p})$ and $I_p(\mathbf{p})$ then introduce, respectively, the angle-independent and the angle-dependent corrections to the initial atomic orbitals.

Since $\phi_1(\mathbf{p})$ satisfies, by definition, an equation

$$[(p_0^2)_1+p^2]\phi_1(\mathbf{p}) = \frac{Z'}{\pi^2}\int\frac{d\mathbf{p}'}{|\mathbf{p}-\mathbf{p}'|^2}\,\phi_1(\mathbf{p}'), \qquad \ldots\ldots(15)$$

where Z' is an effective nuclear charge and $(p_0^2)_1$ the associated value of p_0^2, the dominant term in (12) is

$$\frac{2\pi^2}{Z'}\cos(\mathbf{p}\,.\,\mathbf{R})\,\frac{[(p_0^2)_1+p^2]}{(p_0^2+p^2)}\,\phi_1(\mathbf{p}). \qquad \ldots\ldots(16)$$

In the present case, guided by the screening constant calculation in coordinate space, we shall take $Z'=1{\cdot}228$. The corresponding exact solution of (15) for the 1*s* state is

$$\phi_1(\mathbf{p})=\frac{1}{(p^2+a^2)^2},\quad \text{with}\quad a=(p_0)_1=Z', \qquad \ldots\ldots(17)$$

(16) may therefore be written

$$\frac{2\pi^2}{a}\cos(\mathbf{p}\,.\,\mathbf{R})\,\frac{1}{(p^2+a^2)(p^2+b^2)},\qquad (b=p_0).$$

Comparison with (11) shows that the principal modification introduced by iteration is the replacement of the simple 1*s* atomic orbital by the function

$$\phi_1'(\mathbf{p})=\frac{1}{(p^2+a^2)(p^2+b^2)}. \qquad \ldots\ldots(18)$$

In coordinate space a more satisfactory molecular orbital might therefore be a superposition of atomic orbitals such as

$$\frac{\exp(-ar)-\exp(-br)}{r}\qquad\left(\text{or}\quad \frac{2\sinh\frac{1}{2}(b-a)r}{r}\,.\exp\{-\tfrac{1}{2}(a+b)r\}\right),$$

this being the Fourier transform of (18), rather than of the simple 1*s* orbitals, $\exp(-ar)$.

Since (18) is not appreciably more complicated than (17) we shall accommodate this result in our treatment at the outset by employing $\phi_1'(\mathbf{p})$ rather than $\phi_1(\mathbf{p})$ as an initial atomic orbital. A further complete iteration will now be carried through with the inclusion of angle-dependent corrections. As in I we shall take the best available values of the parameters, namely $p_0^2=2{\cdot}20523$ A.U.; $R=1{\cdot}000$ A.U.

(i) *Spherically Symmetrical Correction*

We deal first with the angle-dependent part of the iterated function. With the new initial orbital, $\phi_1'(\mathbf{p})$, it is possible to evaluate $I_s(\mathbf{p})$ by standard methods of the calculus of residues. We find

$$I_s(\mathbf{p})=2\cos(\mathbf{p}\,.\,\mathbf{R})\,\frac{2\pi^2}{(a^2-b^2)p}\tan^{-1}\left\{\frac{(a-b)p}{ab+p^2}\right\}. \qquad \ldots\ldots(19)$$

The numerical magnitudes involved allow us to replace the $\tan^{-1}$ by the first term in its expansion with a maximum error of the order $\frac{1}{3}\%$. Since it would be presumptuous to regard this as significant we take simply

$$I_s(\mathbf{p})=\cos(\mathbf{p}\,.\,\mathbf{R})\,\frac{4\pi^2}{(a+b)(ab+p^2)}. \qquad \ldots\ldots(20)$$

Ultimately, as already noted, the momentum distribution function $P(p)$ is required. From (12)

$$P(p)=\int\frac{p^2d\omega}{(p_0^2+p^2)^2}[I_s^2+2I_sI_p+I_p^2]. \qquad \ldots\ldots(21)$$

The term in I_s^2 arises, as we have seen, from the partially corrected initial function; the second term in the integrand introduces a further correction whose order of magnitude justifies the neglect of the remaining, second order, term. We are here concerned with the angle-independent term: a simple integration gives

$$P_s(p) = \int_0^{2\pi}\int_0^{\pi} \frac{p^2[I_s(\mathbf{p})]^2}{(p_0^2+p^2)^2} \sin\theta\, d\theta\, d\phi = \frac{2\pi p^2}{(p_0^2+p^2)^2}\left[\frac{4\pi^2}{(a+b)(ab+p^2)}\right]^2\left[\frac{\sin 2pR}{2pR}+1\right]. \qquad \ldots\ldots(22)$$

This curve is compared with the screening constant approximation in Figure 1, where it is scaled to the same peak height.

(ii) *Polarization Correction*

Before the correction due to polarization of the initial atomic orbitals can be computed it is necessary to obtain a series expansion of the integral $I_p(\mathbf{p})$ which cannot be evaluated in closed form. Writing (14) as

$$I_p(\mathbf{p}) = 2\,\mathrm{Re}\exp(i\mathbf{p}\,.\,\mathbf{R})\,.\,I(\mathbf{p}),$$

we may take polar coordinates about $\mathbf{p}$ as axis and, with $\mathbf{R}=(R,\theta,0)$, $\mathbf{p}'=(p',\theta',\phi')$, write

$$I(\mathbf{p}) = \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \frac{\exp\{-2ip'R(\sin\theta\sin\theta'\cos\phi'+\cos\theta\cos\theta')\}p'^2\,dp'\sin\theta'\,d\theta'\,d\phi'}{(p^2+p'^2-2pp'\cos\theta')(p'^2+a^2)(p'^2+b^2)}.$$

The ϕ' integration is easily accomplished (Sommerfeld's Integral), giving

$$I(\mathbf{p}) = 2\pi\int_0^{\pi}\int_0^{\infty} \frac{\exp\{-2ip'R\cos\theta'\cos\theta\}\,.\,J_0(2p'R\sin\theta'\sin\theta)p'^2dp'\sin\theta' d\theta'}{(p^2+p'^2-2pp'\cos\theta')(p'^2+a^2)(p'^2+b^2)}. \qquad \ldots\ldots(23)$$

The remaining angle integration is awkward but may be performed term by term on making the expansion

$$\frac{1}{(p^2+p'^2-2pp'\cos\theta')} = \frac{1}{(p^2+p'^2)}\{P_0(\cos\theta')+\mu_1P_1(\cos\theta')+\ldots.\}, \qquad \ldots\ldots(24)$$

which is obtained from a binomial expansion by expressing the powers of $\cos\theta'$ in terms of the Legendre functions. The μ's are, of course, functions of p and p'.

Subsequently the following result, due to Gegenbauer (1874) may be employed

$$\int_0^{\pi} \exp(iz\cos\theta\cos\psi)\,.\,J_{\nu-\frac{1}{2}}(z\sin\theta\sin\psi)\,.\,C_r^{\nu}(\cos\theta)\,.\,\sin^{\nu+\frac{1}{2}}\theta\,d\theta$$
$$= \left(\frac{2\pi}{z}\right)^{\frac{1}{2}}(i)^r\sin^{\nu-\frac{1}{2}}\psi\,.\,C_r^{\nu}(\cos\psi)\,.\,J_{\nu+r}(z).$$

Assuming the validity of (24) and remembering $P_r(\cos\theta') = C_r^{\frac{1}{2}}(\cos\theta')$, an application of this result with $\nu=\frac{1}{2}$ gives finally

$$I(\mathbf{p}) = 4\pi\int_0^{\infty}\left(\frac{\pi}{p'R}\right)^{\frac{1}{2}}\frac{[A\cos(\mathbf{p}\,.\,\mathbf{R})-B\sin(\mathbf{p}\,.\,\mathbf{R})]p'^2dp'}{(p'^2+a^2)(p'^2+b^2)(p'^2+p^2)}, \qquad \ldots\ldots(25)$$

where

$$A = P_0(\cos\theta)\,.\,J_{\frac{1}{2}}(2p'R) - \mu_2P_2(\cos\theta)\,.\,J_{5/2}(2p'R) + \mu_4P_4(\cos\theta)\,.\,J_{9/2}(2p'R) - \ldots.$$

and

$$B = -\mu_1P_1(\cos\theta)\,.\,J_{3/2}(2p'R) + \mu_3P_3(\cos\theta)\,.\,J_{7/2}(2p'R)$$
$$-\mu_5P_5(\cos\theta)\,.\,J_{11/2}(2p'R) + \ldots.$$

Expressing the μ's explicitly as functions of p and p', and the Bessel functions in terms of elementary functions we might complete the integration expressed in (25). Since we require ultimately $P(p)$ this would not be a very profitable course; instead we compute directly

$$P_p(p)=\frac{2p^2}{(p_0^2+p^2)^2}\int_0^{2\pi}\int_0^{\pi} I_s(\mathbf{p})I_p(\mathbf{p})\sin\theta\,d\theta\,d\phi,$$

with **R** as polar axis. The ϕ integration is trivial on account of axial symmetry and we find finally

$$P_p(p)=\frac{16\pi^2p^2}{(p_0^2+p^2)^2}\left[\frac{4\pi^2}{(a+b)(ab+p^2)}\right][C(p)-D(p)], \qquad \ldots\ldots(26)$$

where

$$C(p)=\Sigma_n C_{\theta n}C_{Rn}, \qquad n \text{ even},$$
$$D(p)=\Sigma_n C_{\theta n}C_{Rn}, \qquad n \text{ odd},$$

and

$$C_{\theta n}=\int_{-1}^{+1}\cos^2(pRx)\,.\,P_n(x)\,dx, \qquad n \text{ even},$$
$$=\int_{-1}^{+1}\cos(pRx)\sin(pRx)\,.\,P_n(x)\,dx, \qquad n \text{ odd},$$
$$C_{Rn}=(-1)^{n/2}\int_0^{\infty}\left(\frac{\pi}{p'R}\right)^{\frac{1}{2}}\frac{\mu_n(p')\,.\,J_{n+\frac{1}{2}}(2p'R)p'^2\,dp'}{(p'^2+a^2)(p'^2+b^2)(p'^2+p^2)}, \qquad n \text{ even},$$
$$=(-1)^{(n+1)/2}\int_0^{\infty}\left(\frac{\pi}{p'R}\right)^{\frac{1}{2}}\frac{\mu_n(p')J_{n+\frac{1}{2}}(2p'R)p'^2\,dp'}{(p'^2+a^2)(p'^2+b^2)(p'^2+p^2)}, \qquad n \text{ odd}.$$

The functions $C_{\theta n}$ are most easily built up from

$$I_{\theta n}=\int_0^1\cos(2pRx)\,.\,x^n\,dx, \quad n \text{ even}, \quad =\int_0^1\sin(2pRx)\,.\,x^n\,dx, \quad n \text{ odd},$$

which may all be derived from $I_{\theta 0}$ by the recurrence relations

$$(2pR)I_{\theta,\,n+1}=(n+1)I_{\theta n}-\cos(2pR), \qquad n \text{ even},$$
$$(2pR)I_{\theta,\,n+1}=\sin(2pR)-(n+1)I_{\theta,\,n}, \qquad n \text{ odd}.$$

The integrals C_{Rn} involve the μ's: these are easily obtained explicitly by a power series expansion of $(p^2+p'^2-2pp'\cos\theta')^{-1}$ in $a\cos\theta'$ where $a=2pp'/(p^2+p'^2)(\leqslant 1)$. The cosines may be written in terms of the Legendre functions by recursion. All the integrals are ultimately reducible to sums of integrals of the type

$$\left.\begin{aligned} I(rS/s,p)&=\int_{-\infty}^{+\infty}\frac{p'^r\sin(2p'R)\,dp'}{(p'^2+a^2)(p'^2+b^2)(p'^2+p^2)^s}, \qquad r \text{ odd},\\ I(rC/s,p)&=\int_{-\infty}^{+\infty}\frac{p'^r\cos(2p'R)\,dp'}{(p'^2+a^2)(p'^2+b^2)(p'^2+p^2)^s}, \qquad r \text{ even}.\end{aligned}\right\} \qquad \ldots\ldots(27)$$

Since a rather large number of these integrals must be computed as functions of p it is gratifying to find that for values of $s>2$ numerical differentiation of $I(rS/2)$, $I(rC/2)$ affords a convenient method of tabulation: higher integrals rapidly become smaller and this treatment gives adequate accuracy. The basic integrals $I(rS/2)$, $I(rC/2)$ may be evaluated by calculus of residues and have been tabulated for $r=1, 2, \ldots. 5$. A five-point rule proves adequate for subsequent differentiation.

The results of the complete calculation appear in Figure 1. Comparison of the final curve with that obtained by Duncanson (1941) by transformation of the coordinate space function due to James (1935) shows that the trend towards regions of higher momentum is maintained with increasing accuracy of the wave function.

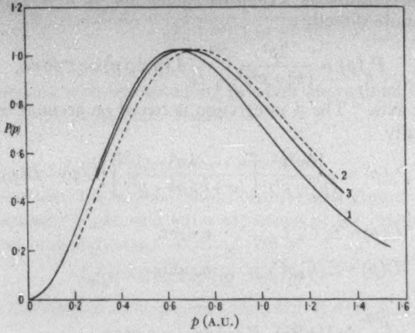


Figure 1. 1. Initial approximation $P^0(p)$. 2. Partially corrected function $P_S(p)$. 3. Corrected function $P(p)$.

§4. EXTENSION TO THE NORMAL MOLECULE

In view of the preceding results it seems likely that even in the case of the normal molecule a preliminary correction of the initial function might be achieved by taking account of the angle-independent distortion of the atomic orbitals. Some progress along these lines has been made and will now be indicated very briefly.

The eigenvalue equation for the normal molecule is found to be

$$(p_0^2+p_1^2+p_2^2)\Phi(\mathbf{p}_1\mathbf{p}_2)=\frac{1}{\pi^2}\left[2Z\left\{\int\frac{\cos\overline{\mathbf{p}_1-\mathbf{p}_1'}\,.\,\mathbf{R}}{|\mathbf{p}_1-\mathbf{p}_1'|^2}\,d\mathbf{p}_1'\Phi(\mathbf{p}_1'\mathbf{p}_2)\right.\right.$$
$$\left.\left.+\int\frac{\cos\overline{\mathbf{p}_2-\mathbf{p}_2'}\,.\,\mathbf{R}}{|\mathbf{p}_2-\mathbf{p}_2'|}\,d\mathbf{p}_2'\Phi(\mathbf{p}_1\mathbf{p}_2')\right\}-\int\frac{d\mathbf{p}}{|\mathbf{p}|^2}\,\Phi(\overline{\mathbf{p}_1-\mathbf{p}},\overline{\mathbf{p}_2+\mathbf{p}})\right]. \qquad \text{......(28)}$$

As in coordinate space, two types of initial function suggest themselves. Taking the product of two molecular orbitals such as

$$\phi^0(\mathbf{p})=\exp(i\mathbf{p}\,.\,\mathbf{R})\,.\,\phi_1(\mathbf{p})+\exp(-i\mathbf{p}\,.\,\mathbf{R})\,.\,\phi_1(\mathbf{p}),$$

and omitting terms corresponding to ionic states gives the "electron-pair" function

$$\Phi^0(\mathbf{p}_1\mathbf{p}_2)=2\phi_1(\mathbf{p}_1)\phi_1(\mathbf{p}_2)\cos\overline{\mathbf{p}_1-\mathbf{p}_2}\,.\,\mathbf{R}. \qquad \text{......(29)}$$

This function is known to be slightly superior to the complete product function and will be made the starting point of our treatment. The momentum distribution corresponding to (29) is easily calculated (Coulson 1941). With hydrogen-like 1*s* functions (17) we adopt Wang's value of the effective nuclear charge

$Z' = a = 1{\cdot}166$ but accept more accurate values of R and p_0^2, namely $R = 0{\cdot}7000$ A.U., $p_0^2 = 3{\cdot}7774$ A.U., since we wish to compute the wave function for the known configuration of the system.

The iterated function $\Phi^1(\mathbf{p}_1\mathbf{p}_2)$ may be written

$$\Phi^1(\mathbf{p}_1\mathbf{p}_2) = \frac{1}{(p_0^2 + p_1^2 + p_2^2)}\left[\{I_1^s\phi_1(\mathbf{p}_2) + I_2^s\phi_1(\mathbf{p}_1)\} + \{I_1^p\phi_1(\mathbf{p}_2) + I_2^p\phi_1(\mathbf{p}_1)\} - I_{12}\right], \qquad \ldots\ldots(30)$$

where

$$I_1^s = 2\cos\overline{\mathbf{p}_1 - \mathbf{p}_2}\,.\,\mathbf{R}\int\frac{d\mathbf{p}_1'}{|\mathbf{p}_1 - \mathbf{p}_1'|^2}\,\phi_1(\mathbf{p}_1')$$

$$I_1^p = 2\,\mathrm{Re}\,[\exp(i\overline{\mathbf{p}_1 + \mathbf{p}_2}\,.\,\mathbf{R})]\int\frac{\exp(-2i\mathbf{p}_1'\,.\,\mathbf{R})}{|\mathbf{p}_1 - \mathbf{p}_1'|^2}\,d\mathbf{p}_1'\phi_1(\mathbf{p}_1'),$$

interchange of suffixes yielding I_2^s, I_2^p, and

$$I_{12} = 2\,\mathrm{Re}\,[\exp(i\overline{\mathbf{p}_1 - \mathbf{p}_2}\,.\,\mathbf{R})]\int\frac{\exp(-2i\mathbf{p}\,.\,\mathbf{R})}{|\mathbf{p}|^2}\,d\mathbf{p}\,.\,\phi_1(\overline{\mathbf{p}_1 - \mathbf{p}})\,.\,\phi_1(\overline{\mathbf{p}_2 + \mathbf{p}}).$$

To compute $P(p_1)$ it is necessary to square the expression (30), integrate over all $\mathbf{p}_2$-space and finally over all directions in $\mathbf{p}_1$-space. Denoting the three terms in (30) by E^s, E^p and E_{12}, and remembering E^p and E_{12} are small compared with E^s it is sufficient to compute

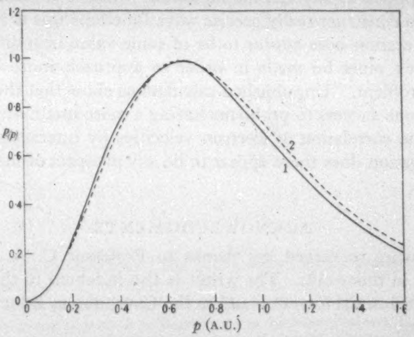


Figure 2. 1. Initial approximation $P^0(p)$. 2. Partially corrected function $P_s(p)$.

$$P(p_1) = \int\!\!\int\frac{d\mathbf{p}_2 p_1^2 d\omega_1}{(p_0^2 + p_1^2 + p_2^2)^2}\,[E^{s2} + 2E^sE^p - 2E^sE_{12}]. \qquad \ldots\ldots(31)$$

The second and final terms in the integrand contain finer details of the corrections due to polarization and to electron correlation respectively: it is, however, noteworthy that the first and dominant term, in itself, is parametrically related to these corrections through p_0^2. For the present we shall evaluate only the major contribution to the momentum distribution, retaining the first term of the integrand alone. In terms of the integrals

$$I(1S/r2, p) = \int_{-\infty}^{+\infty}\frac{p'\sin 2p'R\,dp'}{(p'^2 + a^2)^r(p'^2 + b^2)^2}, \qquad I(2/r2, p) = \int_{-\infty}^{+\infty}\frac{p'^2\,dp'}{(p'^2 + a^2)^r(p'^2 + b^2)^2},$$

where $b^2 = p_0^2 + p^2$, a closed analytical expression for $P(p_1)$ is obtained:

$$P(p_1) = \frac{8\pi^5 p_1^2}{a^2}\left[\frac{\sin(2p_1R)}{2R\,.\,(2p_1R)}\left\{\frac{I(1S/42, p_1)}{(p_1^2+a^2)^2} + \frac{2I(1S/32, p_1)}{(p_1^2+a^2)^3} + \frac{I(1S/22, p_1)}{(p_1^2+a^2)^4}\right\}\right.$$
$$\left. + \left\{\frac{I(2/42, p_1)}{(p_1^2+a^2)^2} + \frac{2I(2/32, p_1)}{(p_1^2+a^2)^3} + \frac{I(2/22, p_1)}{(p_1^2+a^2)^4}\right\}\right].$$

The integrals have been tabulated as functions of p_1 and the result of this first correction of the momentum distribution appears in Figure 2. The broadening of the curve is not so marked as in the cases of the ion: that is, however, a natural consequence of electron correlation since the mutual repulsion of the electrons tends to localize each of them in the vicinity of its own nucleus, making the molecular orbitals of the last section a less appropriate description of the behaviour of each electron.

A more complete treatment of the problem would involve a rather comprehensive tabulation of integrals of the form (14), and is perhaps not easily justified.

§ 5. CONCLUSION

The analysis in this paper makes it clear that the method of iteration is not very conveniently applied to any specific molecule: when it is remembered that our aim has been to construct really precise wave functions this is hardly surprising. Nevertheless iteration does appear to be of some value in indicating the *type* of correction which must be made in order to approach more closely the exact solution of a problem. Unpublished calculations show that the method may be applied with some success to problems having a more qualitative significance, in particular to the correlation of electron velocities by interaction. Only in this type of investigation does there appear to be any prospect of making any further useful progress.

ACKNOWLEDGMENTS

It is a pleasure to record my thanks to Professor C. A. Coulson for his encouragement in this work. The writer is also indebted to the Department of Scientific and Industrial Research and to the Covenantors Educational Trust for Research Grants.

REFERENCES

Coulson, C. A., 1941, *Proc. Camb. Phil. Soc.*, **37**, 55.
Duncanson, W. E., 1941, *Proc. Camb. Phil. Soc.*, **37**, 397.
Duncanson, W. E., and Coulson, C. A., 1941, *Proc. Camb. Phil. Soc.*, **37**, 406.
Gegenbauer, L., 1877, *Wiener Sitzungsberichte*, **75**, 221; see also Watson, G. N., *Theory of Bessel Functions* (Cambridge, 1922).
James, H. M., 1935, *J. Chem. Phys.*, **3**, 9.

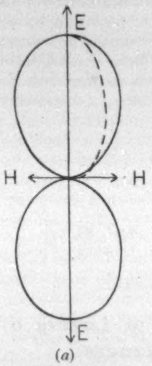
LETTERS TO THE EDITOR

Electrons Ejected by Polarized Radiation

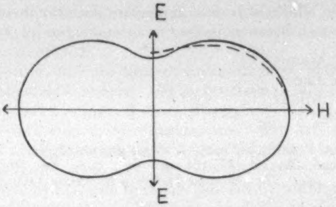
It is well known that, theoretically, the azimuthal angular distribution of photoelectrons ejected by plane polarized radiation follows the "cosine-squared" law (Auger and Perrin 1927). The number of electrons ejected with components between χ and $\chi+d\chi$ is

$$N(\chi)\, d\chi \sim \cos^2 \chi\, d\chi, \qquad \ldots\ldots (1)$$

where χ is the azimuthal angle with respect to the electric vector **E**. (χ is the component of angle measured in the plane perpendicular to the original photon.) The electron intensity shown in polar form in Figure (*a*) is thus greatest in the direction of **E**. Kirchner's mean experimental results (Kirchner 1926, 1927), shown by the dotted curve, are in close agreement with theory.



(*a*)



(*b*)

Comparatively little attention has been given to the azimuthal distribution of recoil electrons ejected by plane polarized radiation. It is, however, known that in this case the emission is greatest in the longitudinal plane containing the magnetic vector **H** ($\chi=\pi/2$) and not in that containing **E**. The reason for this is that the scattered radiation intensity is greatest for $\chi=\pi/2$ or $3\pi/2$, and on Compton's theory it is assumed that momentum is conserved in a collision between an incident photon and a free electron. The azimuthal angles of a recoil electron and scattered photon thus always differ by π, whatever their longitudinal angles may be. If the further assumption is made, that to each scattered photon there corresponds one, and only one, recoil electron, the azimuthal distributions of recoil electrons and scattered photons will be identical, although of course the longitudinal distributions will be quite different.

According to the Klein–Nishina theory (Heitler 1944) the number of photons scattered at a longitudinal angle θ into the element $d\Omega$, and with a given polarization, is

$$N(\theta)_X d\Omega \sim \left(\frac{\nu_\theta}{\nu_0}\right)^2 \left(\frac{\nu_\theta}{\nu_0}+\frac{\nu_0}{\nu_\theta}-2+4\cos^2 X\right) d\Omega, \qquad \ldots\ldots (2)$$

where X is the angle between the electric vectors of the incident and scattered radiations. Following Heitler (1944), consider the scattered radiation as composed of two components, with polarizations **E**′ at right angles. For the first, **E**′ is perpendicular to **E**, so that cos X = O. For the second, **E**′ is in a plane parallel to that containing the direction of the primary photon and its electric vector **E**. In this case, $\cos^2 X = 1 - \cos^2\chi \sin^2\theta$. Taking the mean of these components, the number of photons ejected in the direction (χ, θ) becomes, from (2)

$$N(\chi, \theta)\, d\Omega \sim \left(\frac{\nu_\theta}{\nu_0}\right)^2 \left(\frac{\nu_\theta}{\nu_0} + \frac{\nu_0}{\nu_\theta} - 2\cos^2\chi \sin^2\theta\right) d\Omega. \qquad \ldots\ldots (3)$$

In (2) and (3), $\nu_\theta = \nu_0/[1 + \alpha(1 - \cos\theta)]$, and $\alpha = h\nu_0/mc^2$, the energy of the radiation in terms of mc^2.

The azimuthal distribution of the recoil electrons, which is the same as that of the scattered photons, is now obtained by integrating (3) over θ

$$n(\chi)\, d\chi \sim d\chi \int_0^\pi \left(\frac{\nu_\theta}{\nu_0}\right)^2 \left(\frac{\nu_\theta}{\nu_0} + \frac{\nu_0}{\nu_\theta} - 2\cos^2\chi \sin^2\theta\right) \sin\theta\, d\theta. \qquad \ldots\ldots (4)$$

This can be put in the form

$$n(\chi)\, d\chi \sim (1 - A\cos^2\chi)\, d\chi. \qquad \ldots\ldots (5)$$

The value of A in (5) can be determined analytically from (4) for the case of $\alpha = 0$ $(\nu_\theta = \nu_0)$. Its value is then 2/3. This means that the ratio of the electron emission in the direction **H** to that in the direction **E** is 3 to 1. Graphical integration for other cases shows that, as α increases, the value of A diminishes. For $\alpha = 1$, $A = 0.52$; for $\alpha = 10$, $A = 0.16$. At very high energies the electrons are emitted with azimuthal symmetry by *polarized* radiation. The polar distribution is shown in Figure (*b*) for $\alpha = 0.274$ (140 kev.). The dotted curve, which is in close agreement with the theoretical, represents Kirchner's mean results for recoil electrons ejected in air and argon by radiation of this energy.

17 Parkfield Avenue,
Headstone Lane,
Harrow, Middx.
8th June 1949.

K. H. SPRING.

AUGER, P., and PERRIN, F., 1927, *J. Phys. Radium*, **8**, 93.
KIRCHNER, 1926, *Phys. Z.*, **27**, 385 and 799 ; 1927, *Ann. Phys., Lpz.*, **83**, 521.
HEITLER, W., 1944, *The Quantum Theory of Radiation* (Oxford : University Press), p. 154 (equation 47 *et seq.*).

Theory of Photoconductivity of Layers of Semiconducting Substances

In a previous paper (Schwarz 1948) it has been suggested that adsorption of oxygen ions during the evaporation of semiconducting substances and after the deposit has been formed is an essential condition for the production of photoconductive layers. Based on the experimental results, a qualitative theory of the photoconductivity of these substances is proposed.

During the evaporation of the substance in an electric discharge, several types of adsorption forces are probably active : (i) Van der Waals forces. (ii) Electrostatic forces between the positively charged evaporating particles and the negatively charged oxygen ions. Both kinds of ions are produced by electron bombardment in the electric discharge. (iii) Forces arising from the surface energy of the evaporating particles. (iv) Forces involving valency due to unsaturated valencies or to a small excess of the metallic constituent of the compound as a consequence of slight dissociation during the evaporation.

Each particle when deposited is thus surrounded by an insulating layer of oxygen ions which form a coherent film round the semiconductor particles. This intercrystalline layer fulfils three functions : (*a*) it prevents the coagulation of the particles; (*b*) by the formation of a double layer with the negative side away from the particle it constitutes a potential barrier for the exit of electrons from the particle into the neighbouring particles; (*c*) it is the source and carrier of electrons which by thermal excitation and/or radiation are liberated, accelerated by the external field and form a part of the current.

The transport of electrons along a continuous layer of adsorbed oxygen ions may be visualized as hole conductivity ; a doubly charged oxygen ion loses one electron to a neighbouring singly charged oxygen ion, thus becoming singly charged, and in turn takes up another electron.

A cataphoretic migration of the oxygen ions in an applied electric field, especially at higher temperatures, is also possible.

There are many observations on semiconductors which indicate the presence of a surface layer with special properties and there is strong evidence for a conductivity along the surface different from the bulk conductivity (Leo 1931, 1932, Dubar 1938). It is tentatively suggested that the hole conductivity as discussed above is at least partly responsible for the observed surface conductivity and that the surface levels introduced by Tamm (1932) and investigated by others are, in the case of photoconductive layers, occupied by oxygen ions, adsorbed on the internal surfaces of the semiconductor particles.

The magnitude of hole conductivity will depend on the number of holes and on their activation energy, and these in turn depend on the thickness of the adsorbed layer, on the bond strength between the semiconductor particles and the oxygen ions and on the metal excess within the particle, as will be discussed later.

The adsorption of gas on the particles *during* the evaporation, or in general before the bond between the particles can be established, is essential for the formation of a barrier layer. No heat treatment will produce a sensitive layer if the substance is perfectly free from oxygen and if no oxygen is admitted and adsorbed by the substance while the bond between neighbouring particles is interrupted.

For the formation of an effective barrier layer and a photosensitive deposit it is furthermore essential that a small metal excess within the particle is produced either during the evaporation or after formation of the deposit by a suitable heat treatment. If by unsuitable heat treatment the oxygen ions are removed by desorption, diffusion into the lattice or oxide formation, the photoconductivity is destroyed.

The sensitive layer then possesses at the same time an excess of metal and non-metal. This small metal excess fulfils the following functions: (i) By producing a positive charge of the particle it facilitates the adsorption of negatively charged oxygen ions. (ii) It acts as a donor of electrons producing electrical conductivity of the particle. The activation energy will decrease with increasing metal excess and may even become negative. (iii) The work function may decrease with increasing metal excess. (iv) New absorption bands may be produced. Further experiments have to be made to clarify points (iii) and (iv).

The final barrier layer has most likely a complex structure consisting of a chemi-sorbed layer near the surface of the particle and physically adsorbed layers on top of it.

The bond strength between the oxygen ions and the particle decreases with increasing layer thickness so that the ions of the outermost layer approach the state of free oxygen ions for which the electron affinity for the second electron is negative. Although this state is not reached for adsorbed oxygen ions it indicates that the activation energy for the production of holes will decrease with increasing barrier layer thickness. With increasing metal excess and increasing barrier layer thickness the wavelength threshold for the same substance should therefore shift to longer wavelengths.

Under the influence of radiation absorbed in the insulating layer positive holes are created which render this layer conductive and exert attraction on the electrons within the particle. At the same time radiation absorbed by the semiconductor particle causes an increase of the number and the energy of these electrons, thus increasing the pressure on the barrier layer. The potential barrier breaks down, and electrons can pass between the particles, constituting the photocurrent. It follows that any quantum yield greater than unity is possible, and that under illumination the ratio of excess to defect conductivity should increase.

The electrical and optical properties of the photosensitive deposit will consequently depend on the distribution, size and properties of the semiconductor particles, on the properties of the adsorbed layer and on the interaction between them.

It follows from the foregoing that the photoconductive layer represents a state of non-equilibrium for the ionic and electronic constituents and may be considered as a disperse system. The similarity in colour for layers of different substances and the kind of colour suggest that the size of the particles is fairly uniform for different substances and within the colloidal range.

If adsorption is the essential factor in the production of these films one should expect to find these layers catalytically active. The high catalytic activity of many sulphides and oxides is well known, and the activating treatment of these catalysts closely resembles the heat treatment of the photoconductive layers aiming at the appearance of a new phase—the small metal excess in the case of the photoconductive layer—without making it the predominant component (Hütting 1935). Rectifier plates of Cu_2O showed catalytic activity (Ostwald and Erbring 1931) and the efficiency of rectification increased with the activity, indicating that adsorption of oxygen ions may be essential for rectification of Cu_2O plates.

My thanks are due to Hilger and Watts Ltd. for permission to publish this letter.

Hilger and Watts Ltd.,
98 St. Pancras Way,
Camden Road,
London, N.W. 1.
22nd June 1949.

E. Schwarz.

Dubar, L., 1938, *Ann. Phys., Paris*, **9,** 5.
Hüttig, G. F., 1935, *Z. Elektrochem.*, **41,** 527.
Leo, W., 1932, *Ann. Phys., Lpz.*, **9,** 347 ; 1932, *Ibid.*, **15,** 129.
Ostwald, W., and Erbring, H., 1931, *Colloid Z.*, **57,** 7.
Schwarz, E., 1948, *Nature, Lond.*, **162,** 614.
Tamm, T., 1932, *Phys. Z. Sowjet*, **1,** 733.

REVIEWS OF BOOKS

A Text Book of Heat, by LeRoy D. Weld. 1st Edition. Pp. x+436. (New York: The Macmillan Co., 1948). 25*s*.

This is described as a book for upper classmen, and has been based on the material used for 35 years as lecture notes in a junior-senior course in heat. This does not mean that it is old-fashioned, for modern work has been incorporated wherever it is needed. The properties of liquid helium receive due attention, the totally enclosed absorption type of refrigerator is described, and the International Temperature Scale properly treated.

The general outlook of the book is good. It aims evidently at presenting the student with principles rather than with practice, and many teachers may feel that it gives too little of experimental details; whether he is telling the student how to measure heat quantities or temperatures, or describing the evidence which led to the acceptance of the first law of thermodynamics, the author gives clear explanations of the principles of the experiment, but very little about the practical dodges adopted or required.

The scope of the book includes thermodynamics and statistical mechanics, but the latter suffers because quantum theory is always treated as something outside, to be superimposed on the classical theory. Apart from this, it is an excellent treatment.

The chapters are all accompanied with well selected sets of exercises, and with references for further reading. These show a distinctly American bias, which is, indeed, evident throughout the book. The attribution of the theory of light darts to Epstein may be an example of this, or may be a misprint.

As a whole, the book is very sound, and may give some very useful ideas and hints to those engaged in teaching this subject at the intermediate or pass degree level. J. H. A.

The Theory of Solutions of High Polymers, by A. R. Miller. Pp. vii+118. (Oxford: University Press, 1948). 12*s*. 6*d*.

This book summarizes, in a short space, modern ideas on the properties and behaviour of long molecules in solution. It can be said immediately that, so far as chemists are concerned, only those with an unusual flair for the mathematical approach will find the subject matter readily comprehensible. This is by no means the fault of the author's presentation, which is extremely good, but is due solely to the method of approach necessary in dealing with a difficult subject. Physicists and mathematicians cannot do otherwise than benefit greatly from having recent work brought together in this volume.

The author discusses first the classical theories of solution developed by van't Hoff, van Laer, Gibbs and Lewis, and outlines the assumptions necessary for the derivation of such expressions as Raoult's Law. He then explains the nature of an eigenfunction and of the grand partition function, and their application, first to general cases and then to more specific solutions.

The second chapter is concerned with a description of long chain molecules and the connection between structure and configurational energy. Consideration is then given to the evaluation of the " combinatory factor ", using Bethe's method and modifications thereof. The method involves a statistical examination of sites occupied by molecules in a quasi-crystalline array. It is treated in detail for dimer and trimer molecules. There follows a discussion of free energy of mixing of polymer molecules and the derivation of general formulae.

The section having perhaps greatest interest, at any rate from a practical standpoint, is that which compares theory with experimental findings, e.g. vapour pressure, molar free energies of mixing and heats of dilution of such materials as polystyrene and natural rubber. Because of the nature of the initial assumptions, absolute agreement is not to be expected, but it is interesting that theory and practice are sufficiently related to justify confidence in the postulated ideas. Even for mixtures having non-zero energy of mixing, the relationships are sufficiently close, although the general impression given is that much planned experimental work remains to be done.

The last chapter provides illuminating criticisms of the theory, and shows how the relative sizes of the polymer segments and solvent molecules, and the self-coiling of polymers may affect results. Comments are also made on the validity of the quasi-crystalline model adopted, and it is suggested that further advances may be made by devising a somewhat more satisfactory physical model.

As stated above, the mathematical character of the book will make reading difficult for all except specialists in the field. To these, the matter summarized can only be of the greatest value; other workers interested in high polymers generally will be stimulated by the new concepts discussed by the author. The style is logical throughout, the format is of the highest quality, and the absence of minor errors provides an excellent example of the care taken by the author in compiling this work. N. J. L. MEGSON.

Radiology Physics, by JOHN K. ROBERTSON. Pp. xvii + 323. Second Edition. (London: Macmillan, 1948.) 22*s*.

During the last few years Professor Robertson's *Radiology Physics* has been widely recognized and used as a valuable introductory course for medical and pre-medical students and for those studying the physical principles underlying Medical Radiology.

The first edition, published in 1940, has been amplified very effectively by the addition of sections on Transmutation of matter and neutrons, Artificial radioactivity, mass and energy, and Uranium fission and atomic piles. In addition, the opportunity has been taken to meet previous criticism by the inclusion of a good deal of further information on ultra-violet and infra-red radiations and their applications in medicine.

In accordance with the trend of events Professor Robertson has also included some discussion of the absolute energy systems of dosimetry, which seem destined to play a more important rôle in the correlation of physical phenomena with biological implications. The book does, therefore, provide an admirable introduction to the more recent developments of pure physics and their applications to Medical Radiology, and will continue to be of great service to those commencing a study of the subject. W. V. M.

X-Ray Optics, by A. J. C. WILSON. Pp. vii + 127. Methuen's Monographs on Physical Subjects. 1st Edition. (London: Methuen, 1949.) 6*s*.

The subject matter of this monograph is best indicated by its subtitle, " The diffraction of x-rays by finite and imperfect crystals ". An opening chapter deals with crystal structure and the Laue diffraction theory and is followed by a discussion of the reciprocal lattice. This is used in the main part of the book, which is devoted to the structure of the diffraction

pattern from crystals with various types of imperfections, e.g. layer lattices with planes of a " wrong " type, alloys in a " disordered " state, distorted crystals.

The mathematical treatment is kept as simple as possible. The complex form for a wave is avoided throughout most of the book, and the reciprocal lattice is treated without vector algebra. It seems to the reviewer that the author is unduly pessimistic as to the mathematical equipment of those physics students who are likely to want to read his book. It is, however, a welcome sign that the need for the reciprocal lattice has at last been conceded by English writers on x-ray diffraction. It is to be hoped that the use of vectors will follow in the not too distant future.

A somewhat disturbing feature in a lively and otherwise well written book is the curious terminology employed. No one will deny that physics is more important than phraseology, but is there really any need for the clumsy, misshapen and misleading expression " three dimensional mistakes ", or, for that matter, for " frozen side bands "—another term favoured by the Cambridge School? M. BLACKMAN.

CORRIGENDUM

"The Establishment of the Absolute Scale of Temperature below 1° K.", by A. H. COOKE (*Proc. Phys. Soc.* A, 1949, **62,** 269).

Line 4, p. 274, *should read*

$$Z_c = 2\{1 + 2\exp(-\delta/\boldsymbol{k}T)\} \quad \textit{instead of}$$

$$Z_c = 2\{2 + \exp(-\delta/\boldsymbol{k}T)\}.$$

Line 32, p. 276, *should read*

$$\chi = \chi_0 \left(1 - \frac{4\pi}{3}\chi_0 + 12\eta\chi_0^2\right)^{-1} \quad \textit{instead of}$$

$$\chi = \chi_0 \left(1 - \frac{4\pi}{3}\chi_0 + 12\eta\chi_0^2\right).$$

CONTENTS FOR SECTION B

PAGE
Mr. C. **Gurney** and Dr. S. **Pearson**. The Effect of the Surrounding Atmosphere on the Delayed Fracture of Glass 469
Mr. R. **Mansfield**. The Electrical Properties of Bismuth Oxide . . . 476
Prof. E. N. **da C. Andrade** and Mr. J. W. **Fox**. The Mechanism of Dilatancy . 483
Mr. A. J. **Kennedy**. The Effect of Instantaneous Pre-Strain on the Character of Creep in Lead Polycrystals 501
Mr. E. W. J. **Mitchell** and Dr. R. W. **Sillars**. Observations of the Electrical Behaviour of Silicon Carbide Contacts 509
Dr. H. G. **Wolfhard** and Dr. W. G. **Parker**. Temperature Measurements of Flames containing Incandescent Particles. 523
Letters to the Editor:
Dr. D. **Walker**. Graphical Representation of a Continuous Noise Spectrum. 530
Contents for Section A. 530
Abstracts for Section A 531

ABSTRACTS FOR SECTION B

The Effect of the Surrounding Atmosphere on the Delayed Fracture of Glass, by C. Gurney and S. Pearson.

ABSTRACT. Round soda-lime-silica glass rods were subjected to a series of given bending moments, and times to fracture were recorded. Experiments were made in air and *in vacuo* of 10^{-1} and 10^{-5} mm. of mercury; and the effects of heating *in vacuo* to drive off absorbed atmospheric constituents were also investigated. For the same prior heat treatment the slopes of the curves of stress against time to fracture decrease with decrease of pressure. At 10^{-5} mm. Hg, after heat treatment, the curve was very flat. It is evident that the main cause of delayed fracture is attack of the glass by atmospheric constituents. Experiments in which glass was subjected to air from which water vapour and carbon dioxide were removed separately and together showed that both these substances cause delayed fracture. Other constituents of the atmosphere seem relatively ineffective. These conclusions are in agreement with the work of Preston and his collaborators. It is further concluded that these atmospheric constituents can be supplied by capillary liquid contained in surface cracks and by migration from surface layers as well as by direct attack from the gaseous phase.

The Electrical Properties of Bismuth Oxide, by R. Mansfield.

ABSTRACT. The conductivity σ and thermo-electric power dE/dT of bismuth oxide is measured over a temperature range of 680° to 150° c., and for variations of oxygen pressures from 76 to 10^{-4} cm. Hg. The results above 340°c. and at atmospheric pressure are well represented by the formulae $\sigma = A_1 \exp(-\alpha/2T)$ and below 340° c. by $\sigma = A_2 \exp(-\alpha/T)$, in which the values of the constants A_1, A_2 and α are very similar for two different specimens. The variation of σ with oxygen pressure above 500° c. and for pressures down to 1 mm. Hg obeys the formula $\sigma = k(P_{O_2})^{\frac{1}{4}}$. The thermo-electric power results may be summarized by the equation $dE/dT = a + b/T$, where a and b are constants, the positive sign of dE/dT indicating that Bi_2O_3 is a defect semiconductor. The variations of dE/dT with oxygen pressure are small, less than 10%, except at pressures less than 10^{-3} cm. Hg, when the sign of dE/dT changes at temperatures greater than 550° c. Under these conditions it is suggested that Bi_2O_3 becomes an excess or intrinsic semiconductor. The results are discussed and good agreement is obtained between theory and experiment if it is assumed that normal Wagner and Schottky lattice defects become of importance at low temperatures when the concentration of free electrons is small.

The Mechanism of Dilatancy, by E. N. DA C. ANDRADE and J. W. FOX.

ABSTRACT. The dilatancy in question is the change of overall volume produced by strain in an assembly of particles. This has been demonstrated in the case of a mixture of sand and water by a photoelectric method of measuring the wetness of the surface. The general phenomenon has been experimentally studied in detail with a two-dimensional hexagonal array of uniform cylinders, part of the free surface of which is loaded with a rigid piston. An alternate increase and decrease of overall volume as the piston descends has been demonstrated and the mechanism of slip by which it comes about has been established. The elastic stress distribution in the array before slipping takes place has been calculated and shown to be a determining factor. The effect of friction between the component units has been studied. Experiments have also been carried out on three-dimensional irregular arrays of carbon shot and of sand.

The Effect of Instantaneous Pre-Strain on the Character of Creep in Lead Polycrystals, by A. J. KENNEDY.

ABSTRACT. The extension-against-time curves of lead wires that have been subjected to rapid strain just before the experiment may be expressed by the Andrade creep equation $l=l_0(1+\beta t^{1/3}) \exp(\kappa t)$, using the same constants as those which satisfy the creep of the metal under the same constant stress, but with the t value associated with β replaced by $(t+t_0)$, where t_0 is a constant for a given experiment, its value increasing with increasing pre-strain. While for pre-strains less than 10% the value of κ is unchanged by the pre-strain, for greater pre-strains the κ-flow is also modified and appears to approach more nearly to a linear variation with stress.

Observations of the Electrical Behaviour of Silicon Carbide Contacts, by E. W. J. MITCHELL and R. W. SILLARS.

ABSTRACT. The electrical behaviour of individual silicon carbide contacts has been examined using probes and high resistance electrometer circuits to measure the potential drop across a contact. Voltage-current curves were obtained and the well-known departures from Ohm's law at the contact were confirmed. Attempts were made to fit these experimental results and other data to the theoretical expressions for tunnelling through an insulating film, a hypothesis which does not appear to fit the facts very satisfactorily. Observations were made of the effect of heating the material to high temperatures *in vacuo*, of the properties of a newly formed surface and of the capacitance across the barrier layer. These indicate that the " barrier " has complicated properties which are not consistent with a simple tunnelling model. An alternative model is suggested.

Temperature Measurements of Flames containing Incandescent Particles, by H. G. WOLFHARD and W. G. PARKER.

ABSTRACT. Colour temperatures measured on photoflash powders are higher than expected from theoretical calculations. Determinations of the true temperature of stationary flames of aluminium flakes suspended in air have been made (*a*) by a line reversal method and (*b*) by measuring the absolute light intensity. The experimental values obtained were very close to the theoretical flame temperature of about 3,000° C., that is, the boiling point of the oxide. An explanation is offered for the difference between the true temperature and colour temperature for flames containing incandescent particles, the reason being the small size of the radiant particles, which introduce optical anomalies. It is concluded therefore to be unprofitable to measure the colour temperature at all as this has no relation to the true temperature. Magnesium behaved very similarly, the true temperature being the melting point of MgO, near 2,800° C., whereas the colour temperature is above 3,900° K. The two-colour method of Hottel and Broughton for temperature measurements in hydrocarbon flames is examined in the light of this experience.